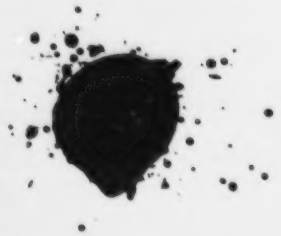
Les PUBLICATIONS DU QUÉBEC

ÉLÉMENTS DE DOCIMOLOGIE

fascicule 4

l'évaluation sommative



Québec

Le contenu de cette publication a été réalisé
par la Direction générale du développement pédagogique
du ministère de l'Éducation

Cette édition a été produite par
Les Publications du Québec
1279, boulevard Charest Ouest
Québec (Québec)
G1N 4K7

BIBLIOTHÈQUE ADMINISTRATIVE
Ministère des communications du Québec
Éléments de catalogage avant publication

Éléments de docimologie: t.4: L'évaluation sommative /[réalisé par la Direction générale du développement pédagogique du Ministère de l'éducation]. — — [Québec]: Le Ministère, 1988.

84 p.
ISBN: 2-551-09002-4

1. Tests et mesures en éducation 2. Classement et notation (Élèves et étudiants) 3. Évaluation en éducation I. Québec (Province). Direction générale du développement pédagogique II. Titre: L'évaluation sommative
E3 D41 E43 / t.4

© Gouvernement du Québec
Dépôt légal — 1er trimestre 1988
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISBN 2-551-08207-2

Sommaire

Introduction

Chapitre premier

Définition

1.1 Qu'est-ce que l'évaluation sommative? 6

1.2 Comment se distingue l'évaluation sommative de l'évaluation formative? . 7

A. La décision 7

B. Le moment 8

C. Les apprentissages évalués. . . . 8

1.3 Résumé 9

1.4 Références 9

Chapitre deuxième

Le contexte de l'évaluation sommative

Préambule 12

2.1 Quels sont les différents éléments qui peuvent influencer l'évaluation sommative? . . . . . 1[illegible]

A. L'approche pédagogique 13

B. Le régime pédagogique 14

C. Les programmes d'études . . . 14

D. Les méthodes d'enseignement . . 15

2.2 Pour qui l'évaluation sommative est-elle importante? 16

A. Les élèves 16

B. Les parents 16

C. Les enseignants 17

D. Les professionnels 17

E. Les administrateurs 17

2.3 Comment partager les responsabilités?. 17

A. Le ministère de l'Éducation . . . 18

B. La commission scolaire 18

C. Le personnel enseignant et professionnel 18

2.4 Résumé 19

2.5 Références 19

Chapitre troisième

Comment planifier l'évaluation sommative?

Préambule 23

3.1 Planifier l'étape décision 25

A. Identifier le type de décision à prendre 25

B. Faire l'inventaire des possibilités . 26

C. Établir les règles à suivre . . . 27

3.2 Planifier l'étape jugement 27

A. Comment juger de la compétence d'un élève en regard d'un programme d'études? 28

B. Comment juger des autres éléments pertinents à l'orientation de l'élève dans sa carrière scolaire? 32

C. Comment traiter les cas limites et les cas surprises? 33

D. Déterminer les moments où l'on portera les jugements 34

3.3 Planifier l'étape mesure 34

A. Choisir la façon d'obtenir les informations nécessaires aux jugements à porter. 36

B. Choisir les façons d'interpréter les informations et d'établir le seuil de réussite s'il y a lieu 39

C. Déterminer les moments où l'on procédera à la mesure 44

3.4 Choisir une façon de conserver les données 45
3.5 Résumé 47
3.6 Références 48

Chapitre quatrième
La mesure en évaluation sommative

Préambule 52
4.1 Quelles qualités rechercher pour une épreuve? 52
A. La validité 52
B. La fidélité 54
4.2 Que faut-il surveiller au moment de la passation de l'épreuve? 55
4.3 Quelles précautions faut-il prendre lors de la correction des réponses? . . . 59
4.4 Comment interpréter les données? . . 60
A. L'analyse d'items 60
B. L'interprétation des données. . . 62
4.5 Quelles qualités assurer quand on prend les informations complémentaires? 64
4.6 Résumé 64
4.7 Références 65

Chapitre cinquième
Le jugement et la décision

Préambule 68
5.1 À quel moment juger de la situation de l'élève? 69
5.2 Comment juger? 69
A. Combiner les multiples données recueillies 69
B. L'erreur de jugement 71
C. Cas limites et cas imprévus . . . 72
D. La répartition des responsabilités . . 72
5.3 Comment prendre la décision? . . . 73
5.4 Comment consigner les jugements et les décisions? 74
5.5 Résumé 76
5.6 Références 76
Conclusion 77

Liste des tableaux

Tableau 1
Caractéristiques de l'évaluation selon la fonction qu'elle exerce 8

Tableau 2
La planification de l'évaluation sommative 24

Tableau 3
Types d'erreurs possibles 25

Tableau 4
Liste de questions pour vérifier la validité d'une épreuve avant de l'utiliser 53

Tableau 5
Types d'indices de fidélité d'une épreuve selon l'approche normative . 56

Tableau 6
Les caractéristiques des scores et des décisions pour les trois types de fidélité selon l'approche critériée 57

Tableau 7
Types de jugements erronés 72

Introduction

L'évaluation des apprentissages est une démarche dont on se sert pour prendre plusieurs décisions dans le système scolaire. L'évaluation faite dans le but d'ajuster les interventions pédagogiques et de faciliter l'apprentissage de l'élève est dite formative; c'est à ce type d'évaluation qu'a été consacré le fascicule 3.

Dans le présent document, on décrit l'évaluation sommative. Depuis longtemps, ce type d'évaluation sert à prendre des décisions qui concernent la sanction des études ou la carrière scolaire de l'élève, comme le passage à la classe supérieure. Jusqu'ici l'établissement des règles de décisions était cependant plus ou moins implicite. Plus encore, les jugements à porter n'étaient pas formulés et la plupart du temps, ils étaient confondus avec les mesures ou les décisions prises. Ici, on distingue les trois étapes de la démarche d'évaluation: mesure, jugement, décision.

Au moment de la planification, chapitre troisième, on identifie la décision à prendre et on prévoit le déroulement des étapes «jugement» et «mesure» en congruence avec les décisions auxquelles elles serviront tout en tenant compte du partage des responsabilités. Par la suite, lorsque l'on présente la réalisation de cette démarche, on explicite d'abord comment prendre les informations pertinentes, chapitre quatrième, puis comment porter les jugements sur lesquels s'appuieront les décisions, chapitre cinquième; l'étape décision constitue alors l'aboutissement et la justification de toute la démarche.

Soulignons que dans la plupart des situations d'évaluation sommative, les règles qu'on établit et les instruments qu'on utilise devraient s'inscrire dans une approche critériée où l'on tente de bien cerner la compétence de l'élève par rapport à un programme donné. C'est pourquoi la présentation faite dans ce fascicule relève surtout d'une telle approche.

Notes préliminaires

Nous ne traitons ici que de l'évaluation sommative des apprentissages des élèves. Les autres objets de l'évaluation sommative (de programme, d'enseignement, d'institution...) ne sont pas touchés dans ce document. Il faut être conscient que dans ces cas il y a des façons de faire particulières qui peuvent être différentes de celle qui est expliquée ici. Il s'agit donc d'une présentation de l'évaluation sommative des apprentissages dans différents contextes. On y précise les conditions qui assureront une démarche de qualité ainsi que des façons de faire pour y arriver.

Définition

1.1 Qu'est-ce que l'évaluation sommative? 6
1.2 Comment se distingue l'évaluation sommative de l'évaluation formative? . 7
A. La décision 7
B. Le moment 8
C. Les apprentissages évalués . . . 8
1.3 Résumé 9
1.4 Références 9

Chapitre premier

Dans ce chapitre, nous verrons que:

1. La pratique de l'évaluation sommative est très ancienne mais le terme est récent.
2. L'évaluation sommative est une démarche qui vise à prendre des décisions reliées à l'orientation de l'élève dans sa carrière scolaire ou à la sanction des études de cet élève.
3. On distingue l'évaluation sommative de l'évaluation formative selon trois aspects: la décision à prendre, le moment où l'on fait l'évaluation et les apprentissages évalués.

1.1 Qu'est-ce que l'évaluation sommative?

Même si le terme «évaluation sommative» est assez récent , sa pratique n'est pas nouvelle.

L'expression «summative evaluation» a été introduite par Scriven en 1967:

> L'évaluation sommative d'un programme... est réalisée après avoir complété ce dernier pour le bénéfice de groupes ou décideurs externes... quoiqu'elle puisse être réalisée par des évaluateurs internes ou externes ou par une combinaison des deux... (Scriven, 1981, p. 150).

C'est aussi pour ce même auteur «le processus de détermination du mérite ou de la valeur de quelque chose ou le produit de ce processus» (Scriven, 1981, p. 53). Le fait de donner un nom à cette réalité qu'est l'évaluation sommative permet de s'approprier et d'analyser de plus en plus ce type d'évaluation. C'est en fait ce que nous ferons tout au cours de ce document.

Ce qui était d'abord prévu par Scriven pour les programmes s'est graduellement étendu à d'autres situations scolaires.

Plus tard Bloom (1971) définissait l'évaluation sommative ainsi:

> C'est le type d'évaluation employée à la fin d'un cours, d'un cycle ou d'un programme dans un but de classification, d'évaluation du progrès ou dans l'intention de vérifier l'efficacité d'un programme ou d'un cours. La caractéristique essentielle de l'évaluation sommative c'est qu'un jugement est porté sur les apprentissages réalisés par les élèves ou sur l'efficacité d'un programme au terme du déroulement de la démarche.

Gilbert De Landsheere (1974) rejoint Bloom lorsqu'il dit que l'évaluation sommative porte «soit sur une partie importante d'un cours, soit même sur le tout. Mais il peut aussi s'agir d'objectifs généraux tels que l'acquisition d'une attitude ou d'une capacité».

On remarque par ailleurs que Weiss (1977) restreint l'évaluation sommative à la fonction certificative et qu'il la fait porter sur les objectifs terminaux.

On note enfin que Tourneur (1983, p. 18) adopte un point de vue semblable à celui de Bloom. Pour lui l'évaluation sommative donne lieu à deux types de décisions: «délivrer un diplôme qui donne accès à une profession ou attester la maîtrise des compétences de 'survie' sociale (certification externe) et vérifier que l'étudiant possède les prérequis nécessaires pour poursuivre sa formation (certification interne).» Selon Tourneur l'évaluation sommative est utile aussi pour des décisions à caractère prédictif.

On peut donc parler d'une évaluation prédictive. Ce type d'évaluation peut porter sur des objets différents et recourt à des méthodes appropriées à son intention: orienter l'élève dans sa carrière scolaire. Cela aurait pu être traité en soi mais, compte tenu de la définition adoptée ci-dessous, l'évaluation prédictive est considérée dans ce document comme un volet de l'évaluation sommative.

En effet, lors de la préparation du cadre général de la présente série et notamment du Lexique, le ministère de l'Éducation du Québec a adopté pour l'évaluation sommative, la définition suivante:

> «Démarche qui vise à porter un jugement sur le degré de réalisation des apprentissages visés dans un programme ou une partie terminale de programme ou encore

d'un ensemble équivalent d'apprentissages faits en dehors de l'école, en se basant sur des données pertinentes, pour prendre des décisions relatives:

— au passage à la classe supérieur,

— à la sanction des études,

— à la reconnaissance des acquis expérientiels (équivalences),

— à l'orientation des élèves.»

Cette définition tient compte du contexte scolaire québécois: programmes d'études, groupe-classe, cheminement particulier, etc.

C'est donc dire que c'est l'évaluation sommative que l'on utilise lorsqu'il s'agit de sanctionner les études de l'élève et de le guider dans sa carrière scolaire. L'évaluation sommative a donc une portée externe, c'est-à-dire orientée vers la société, lorsqu'elle sert à prendre des décisions relatives à la sanction des études et, dans certains cas, à la reconnaissance des acquis expérientiels. L'évaluation sommative a également une portée interne, c'est-à-dire orientée vers le cheminement dans le système scolaire, quand elle sert à prendre des décisions relatives au passage à la classe supérieure et à l'orientation des élèves à moyen terme.

1.2 Comment se distingue l'évaluation sommative de l'évaluation formative?

Qu'elle soit formative ou sommative, l'évaluation est considérée comme une démarche qui se réalise en trois temps. D'abord, on recueille, organise et interprète les informations nécessaires: c'est l'étape **mesure**. Ensuite, on porte un jugement sur la situation, la considérant satisfaisante ou non: c'est l'étape **jugement**. En un troisième temps, on étudie les différents choix possibles et on prend une **décision**.

L'évaluation sommative se distingue de l'évaluation formative selon les aspects suivants:

a) la décision à prendre,

b) le moment où l'évaluation a lieu au cours de l'enseignement/apprentissage,

c) les apprentissages évalués.

L'évaluation sommative tient compte d'un passé plus ou moins récent pour guider des décisions à plus long terme. L'évaluation formative[1] guide des décisions plutôt immédiates comme d'expliquer différemment une notion à un élève qui n'a pas compris en un premier temps. Elle revêt un caractère plutôt pédagogique en favorisant les apprentissages tandis que l'évaluation sommative fait le bilan à un moment déterminé, sanctionne les savoirs, les savoir-faire et les savoir-être acquis.

A. La décision

La décision prise en évaluation sommative a un caractère statique: on arrête dans le temps et on se prononce sur quelque chose qui est supposé être fini. À cause de l'ampleur des décisions prise (passer à la classe supérieure, avoir son diplôme, entreprendre tel cours, etc.), on a moins d'emprise sur les conséquences car il est plus difficile de modifier le cours des actions entreprises, plusieurs personnes étant impliquées. Par ailleurs, l'évaluation formative a un caractère dynamique. En effet, si l'on a permis à tort à un élève d'entreprendre une activité d'apprentissage reliée à un nouvel objectif, il est toujours possible de réviser cette décision. Dans ce cas l'élève n'aura pas perdu beaucoup de temps. Il n'en serait pas de même si on lui avait permis à tort de passer à la classe supérieur. L'évaluation formative guide donc des décisions plutôt immédiates tandis que l'évaluation sommative oriente des décisions à plus long terme.

1 Le fascicule 3 traite en détail de l'évaluation formative.

B. Le moment

L'évaluation sommative s'effectue à la fin d'une période d'enseignement/apprentissage consacrée à une partie terminale de programme ou au programme entier; elle concerne le passé. Si pendant l'année on a porté des jugements de type sommatif à plusieurs occasions, on doit se demander à la fin de l'année si ces jugements sont encore justes. L'évaluation formative, pour sa part, s'applique pendant la période même d'enseignement/apprentissage.Elle porte sur le présent ou sur le passé immédiat et se caractérise par une possibilité de retour sur l'apprentissage.

C. Les apprentissages évalués

L'évaluation sommative privilégie la vérification d'apprentissages importants du programme qui sont généralisables, transférables à plusieurs situations. Pour sa part l'évaluation formative, correspondant à la préoccupation d'assurer l'ajustement continu du processus enseignement/apprentissage, porte sur des productions ou des démarches de l'élève qui sont souvent considérées comme transitoires et modifiables.

Pour distinguer l'évaluation sommative de l'évaluation formative, on est souvent porté à classer les démarches d'évaluation selon le moment où elles se produisent, mais la perspective temporelle n'est pas la caractéristique essentielle qui les distingue. Notons que pour déterminer le type auquel appartient une évaluation, il est fondamental de se baser sur la fonction de cette évaluation et non pas seulement sur le moment où elle se produit.

Le tableau suivant présente différentes caractéristiques de l'évaluation selon ses fonctions formative ou sommative. Ces deux types d'évaluation se distinguent principalement sur trois points: les décisions prises, le moment où le jugement est porté et le type d'apprentissage mesuré. Si l'on voulait caractériser l'évaluation sommative, on pourrait dire que les décisions prises à la fin de la démarche donnent lieu à des actions dont le cours est complexe à modifier. On pourrait dire aussi que ces décisions sont basées sur des jugements portés à la fin de la période consacrée à l'enseignement/apprentissage d'un programme ou d'une partie terminale de celui-ci. On pourrait dire enfin que les infor-

Tableau 1

Caractéristiques de l'évaluation selon la fonction qu'elle exerce

	Fonction		
	Formative	**Sommative**	
Intention	Favoriser la progression de l'apprentissage	Sanctionner les apprentissages, reconnaître les acquis	Orienter l'élève dans son cheminement scolaire
Contenu	Les préalables à un apprentissage, la démarche utilisée par l'élève	Les compétences acquises	Les prédicteurs de réussite
Cadre de référence	Les hiérarchies d'apprentissage, les stratégies d'apprentissage, les apprentissages visés	Les objectifs des programmes d'études	Les caractéristiques de la population cible
Exemple d'instrument	Observation du développement de l'habileté à comprendre un conte à l'aide de questions orales	Examen-synthèse	Prise d'information sur les aptitudes, les intérêts

Inspiré de Chardinet (1973, p. 158)

mations permettant de porter les jugements proviennent surtout de la mesure d'apprentissages généralisables, c'est-à-dire transférables à diverses situations.

1.3 Résumé

Dans ce chapitre, nous avons considéré l'évaluation sommative comme une démarche comprenant trois étapes: la mesure, le jugement et la décision. Au moment de la mesure on prend toutes les informations pertinentes puis on porte un jugement éclairé en fonction de la décision à prendre quant à l'orientation de l'élève dans sa carrière scolaire ou à la sanction des études. L'évaluation sommative joue donc un rôle important dans la carrière scolaire et professionnelle de l'élève. Par la sanction des études et la reconnaissance des acquis expérientiels, elle remplit aussi un rôle dans l'organisation sociale.

1.4 Références

BLOOM, Benjamin S., J. Thomas, HASTINGS, George F. MADAUS, *Handbook on Formative and Summative Evaluation of Student Learnig*, New York, McGraw-Hill, 1971, 923 p.

BLOOM, Benjamin S., George F., MADAUS, HASTING, J. Thomas. *Evaluation to Improve Learning*, New York, McGraw-Hill, 1981, 356 p.

CARDINET, Jean. L'adaptation des tests aux finalités de l'évaluation, *Sciences de l'éducation*, 2 - 3, 1973, p. 148 - 182.

CARDINET, Jean. *Pour apprécier le travail des élèves*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, août 1984, 75 p.

CARDINET, Jean. *Des instruments d'évaluation pour chaque fonction*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1983, 81 p.

DE LANDSHEERE, Gilbert. *Dictionnaire de l'évaluation et de la recherche en éducation*, Paris, Presses universitaires de France, 1979, 338 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Eléments de docimologie: Lexique*, Québec, 1985, 23 p.

MORISSETTE, Dominique. *La mesure et l'évaluation en enseignement*, Sainte-Foy, Les Presses de l'université Laval, 1984, 423 p.

PAQUAY, L. *Méthodologie générale. Fonctions basales de l'évaluation*, Document 7, IC-AFOC-CRAF, 03/10/83, 6 p.

POPHAM, W.J. Setting Performance Standards, in *Modern Educational Measurement* , Englewood Cliffs, N.J., Prentice-Hall, 1981, p. 371-399.

SHEPARD, Lorrie. «Standard Setting Issue and Methods», *Applied Psychological measurement*, Vol. 4, N° 4, automne 1980, p. 447 - 467.

SCRIVEN, Michael. *Evaluation Thesaurus*, 3d ed., Edgepress, 174 p.

THURLER, Monica. *Réflexions sur l'évaluation de l'élève*, Rapport d'atelier 1, Groupe de travail 2, Genève, Directeurs du projet SIPRI, 1983, 48 p.

TOURNEUR, Y. *La certification des compétences. Analyse sur ses fonctions et ses méthodes aux différents moments de la formation*, Actes des Rencontres Évaluation formative et formation des enseignants, Namur, 1983, p. 80 - 93.

WEISS, J. *À la recherche d'objectifs terminaux en mathématique au niveau élémentaire*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, IRDP/R 77.03, Février 1977.

Le contexte de l'évaluation sommative

Préambule 12
2.1 Quels sont les différents éléments qui peuvent influencer l'évaluation sommative? 12
A. L'approche pédagogique 13
B. Le régime pédagogique 14
C. Les programmes d'études 14
D. Les méthodes d'enseignement . . 15
2.2 Pour qui l'évaluation sommative est-elle importante? 16
A. Les élèves 16
B. Les parents 16
C. Les enseignants 17
D. Les professionnels 17
E. Les administrateurs 17
2.3 Comment partager les responsabilités? 17
A. Le ministère de l'Éducation . . . 18
B. La commission scolaire 18
C. Le personnel enseignant et professionnel 18
2.4 Résumé 19
2.5 Références 19

Chapitre deuxième

Dans ce chapitre, nous verrons que:

1. Le contexte dans lequel se déroule l'évaluation sommative peut influencer une ou plusieurs étapes de cette démarche.
2. L'approche pédagogique influence le cheminement scolaire de l'élève. Elle pourrait aussi influencer la sanction des études.
3. Le régime pédagogique détermine globalement les décisions à prendre, les règles à suivre et certains moments d'évaluation au cours de la carrière scolaire.
4. Le programme d'études détermine les objets d'évaluation et, d'une certaine façon, les moments où l'on doit procéder à l'évaluation.
5. La méthode d'enseignement peut influencer la façon de prendre de l'information sur les apprentissages de l'élève.
6. L'évaluation sommative est importante pour l'élève parce qu'elle a des conséquences sur sa carrière scolaire et professionnelle. Elle a de plus un impact sur son concept de soi.
7. Pour bien assumer sa responsabilité envers son enfant, le parent a le besoin et le droit d'être informé du résultat des évaluations.
8. Les résultats d'évaluation sommative permettent à l'enseignante ou à l'enseignant d'analyser son enseignement.
9. L'administratrice ou l'administrateur a besoin de l'évaluation sommative pour procéder à une gestion efficace du système scolaire.
10. La responsabilité de l'évaluation sommative est partagée entre l'enseignante ou l'enseignant, la commission scolaire et le ministère de l'Éducation.

Préambule

Quant on examine les problèmes rencontrés en évaluation sommative et qu'on tente de répondre aux interrogations qui surgissent et de réduire les incohérences, on est souvent amené à considérer les éléments fondamentaux qui influencent l'évaluation sommative, c'est-à-dire les choix qui ont été faits de même que les valeurs qui les justifient.

L'évaluation sommative concerne plusieurs personnes ayant des rôles différents dans l'organisation scolaire et implique diverses instances scolaires. Elle met donc en jeu plusieurs façons de faire de même que différentes valeurs et priorités.

De par sa conception et sa structure, notre système d'éducation impose certains choix. Ceux-ci reposent sur des valeurs établies dans «L'école québécoise: Énoncé de politique et plan d'action». Dans ce chapitre, nous identifions les choix qui sont déjà faits par le système d'éducation et la latitude qui reste notamment en ce qui a trait au partage des responsabilités et aux nombreuses décisions reliées à la réalisation de chacune des étapes de l'évaluation sommative.

2.1 Quels sont les différents éléments qui peuvent influencer l'évaluation sommative?

Le contexte dans lequel s'insère l'évaluation sommative comporte plusieurs éléments importants entre autres, l'approche pédagogique, le régime pédagogique, les programmes d'études et les méthodes d'enseignement. Voyons comment ils peuvent influencer l'évaluation sommative.

L'approche pédagogique concerne l'école et son projet éducatif. Elle définit le genre de relation que les gens ont entre eux globalement, le type de discipline établie, le partage des responsabilités entre la direction, le personnel enseignant, l'élève, ses parents et les professionnels. En adoptant une approche pédagogique, l'école accepte explicitement ou implicitement un ensemble de valeurs.

Le régime pédagogique pour sa part, se situe au niveau de l'organisation des cheminements scolaires possibles pour l'élève. Il fixe certaines conditions comme le nombre de crédits à accumuler pour la sanction de études. Selon les valeurs que le régime pédagogique sous-tend, certaines approches pédagogiques seront applicables tandis que d'autres demanderont à être modifiées avant d'être appliquées.

Les programmes d'études fixent, à partir d'un certain consensus, l'ensemble des apprentissages que l'élève doit réaliser. Le contenu des programmes (buts et objectifs) de même que leur forme impliquent des valeurs et des conceptions de l'homme qui peuvent avoir certaines affinités avec des approches pédagogiques. De plus, les programmes d'études peuvent dans certains cas déterminer les méthodes d'enseignement possibles.

Les méthodes d'enseignement sont plus spécifiques que les approches pédagogiques en ce sens qu'elles ne touchent que l'organisation immédiate de l'apprentisssage. Par exemple, si j'utilise une méthode inductive j'essaie d'amener l'élève à découvrir, à partir de plusieurs observations, un principe ou une règle quelconque. Si je procède par déduction, je donne la règle et l'élève développe son habileté à l'appliquer à des cas variés. Les méthodes s'appuient sur les didactiques des matières. On ne peut dire que telle méthode relève de telle approche, les liens entre approches et méthodes n'étant pas aussi directs que cela. Toutefois, il est logique de penser que certaines affinités peuvent exister. Par exemple, on verrait mal une méthode qui accorde beaucoup de place à l'élève être utilisée dans une approche qui

considère l'enseignant comme étant la seule personne qui distribue les connaissances.

Bref, le contexte dans lequel se fait l'évaluation sommative forme un tout aux éléments interreliés. Reste à savoir l'influence de ce contexte sur l'évaluation.

A. L'approche pédagogique

On retrouve au Québec des tendances éducatives assez variées tant chez les parents, les personnels enseignants, que chez les autorités administratives, tendances qui peuvent se concrétiser dans le projet éducatif de chaque école.

Les finalités de notre système d'éducation, exposées dans «L'école québécoise: Énoncé de politique et plan d'action» (1979), s'inscrivent dans une option humaniste, c'est-à-dire centrée sur la personne et axée sur le savoir-être. De plus, tel qu'indiqué dans «Politique générale d'évaluation pédagogique» (Gouvernement du Québec, 1981), on vise pour tous les élèves l'égalité des chances face à l'école. Par ailleurs, comme ce système fait partie d'une société post-industrielle, il est aussi influencé par les valeurs de celle-ci: efficacité, objectivité, rationalité scientifique,... Depuis une dizaine d'années, on a visé la démocratisation du système scolaire, ce qui implique la possibilité de choisir un projet éducatif. C'est dire qu'au Québec, le système d'éducation accepte la coexistence de plusieurs valeurs, de plusieurs tendances. En contrepartie, l'école doit accepter un minimum de directives et d'uniformité assurant l'existence du système malgré les tiraillements inévitables qui s'ensuivent.

L'évaluation sommative est influencée par l'approche pédagogique. En effet, lorsqu'une école choisit une approche pédagogique, elle opte implicitement ou explicitement pour un type de société qu'elle tente d'actualiser. En fait, l'approche pédagogique relève d'une conception particulière de l'homme et de sa place dans la société. Elle implique certaines valeurs et ce sont ces valeurs qui guident l'intervention en classe, l'organisation scolaire, le choix des objectifs à évaluer, la façon de le faire et l'utilisation des résultats.

Afin d'illustrer l'impact de l'approche pédagogique sur l'évaluation sommative, considérons l'école ouverte qui est fondée sur la participation des personnes qui y vivent (élèves, personnel enseignant et autres). Elle se propose d'optimiser la croissance des élèves. Cela se produit entre autres par l'auto-éducation, c'est-à-dire par le retour à chaque individu et à l'environnement où il vit. L'école ouverte est donc fondée à la fois sur l'individualisation et sur l'intégration à la collectivité (Paré, 1977).

On y pratique une pédagogie de la participation où plus rien ne va de soi, où tout est interdépendant et demande une définition constante. On comprendra facilement qu'on y prône l'évaluation formative. Ainsi, Paré (1977, vol. III: 64) dira que ce n'est pas le portrait du résultat de la course qu'il faut, mais la représentation du déroulement de chaque performance. Dans ce contexte, on cherche à évaluer quatre aspects importants. On se demande:

— Quelle est la qualité des activités éducatives?

— Quelle est la qualité de la programmation réalisée par le maître?

— Quelles attitudes visées par l'école ouverte ont été développées?

— Quels apprentissages stables et transférables ont été faits?

Pour des fins d'évaluation sommative, on considérera les apprentissages stables et transférables développés chez l'élève. Comme dans les classes du primaire le

classement peut être effectué sur des bases variées, on ne se pose pas les mêmes questions au même moment. Mais lorsqu'il s'agit du passage au secondaire, de l'octroi de crédits et de la sanction des études, les règles établies par le ministère de l'Éducation s'appliquent.

Quelle que soit l'approche pédagogique choisie, la principale question ici est de déterminer les données pertinentes sur lesquelles doivent se baser les jugements à faire en vue de la sanction des études ou de l'orientation de l'élève dans sa carrière scolaire. Les approches pédagogiques influencent donc le cheminement de l'élève parce qu'elles offrent des organisations différentes (groupe-classe structuré ou non, autogestion du temps d'apprentissage ou non, etc.). Ainsi, dans certains cas, les parents choisiront de placer l'élève dans une autre école parce qu'ils croient que l'approche pédagogique mise de l'avant par cette école sera plus favorable à leur enfant. Si les diverses approches pédagogiques étaient appliquées intégralement, elles pourraient même influencer la sanction des études. Ainsi, dans l'approche de type école ouverte, c'est l'élève qui détermine le contenu d'apprentissage et évalue lui-même l'importance et la signification des apprentissages qu'il fait. Au Québec, le régime pédagogique et les programmes d'études imposent cependant des limites dans le choix des approches pédagogiques et donc dans la variété des formes d'évaluation.

B. Le régime pédagogique

Les règlements concernant le régime pédagogique et les instructions émises par le ministère de l'Éducation encadrent et influencent l'application de l'évaluation sommative de bien des façons. C'est ainsi qu'en 1981, le ministère de l'Éducation établissait l'intensification des échanges d'information entre l'école et les parents, une plus grande précision des contenus des programmes et la clarification du partage des responsabilités en matière d'évaluation. Actuellement au primaire, c'est la commission scolaire, le personnel de l'école et particulièrement l'enseignante et l'enseignant qui sont responsables des évaluations; pour sa part le Ministère intervient principalement au second cycle du secondaire à des fins de sanction des études. Au primaire, on veut assurer le progrès continu de l'élève. Au secondaire, on accorde les crédits pour chaque programme réussi de la 1er à la 5^{e} année secondaire. De plus, le diplôme n'est accordé qu'à ceux qui ont accumulé un minimum de 130 crédits dont 40 crédits obligatoires dans un éventail diversifié et déterminé de matières. Un des buts de cet ensemble de règlements est d'assurer une formation de base équilibrée. Le régime pédagogique fixe des moments particuliers pour effectuer l'évaluation sommative(passage à la classe supérieurre, fin d'un programme). Il détermine globalement l'objet d'évaluation et établit des principes de décisions (passage obligatoire au secondaire après sept ans d'études, réussite obligatoire de tel cours de français au secondaire, etc.).

C. Les programmes d'études

Il faut bien analyser les programmes avant de planifier l'évaluation sommative car certaines de leurs caractéristiques détermineront des façons de faire.

Les programmes n'ont pas tous la même structure: certains sont cycliques tandis que d'autres sont séquentiels. L'évaluation sommative, arrivant en bout de course, se situe donc à des moments différents pour ces deux types de programmes. Une préoccupation demeure toutefois, celle de laisser le plus possible à l'élève le temps qui lui est nécessaire pour apprendre.

Ainsi, en mathématique, le développement de l'habileté à grouper et regrouper des

objets en base dix est visé de la 1re à la 3e primaire. Deux solutions sont alors possibles: attendre à la fin du cycle pour se prononcer sur l'atteinte de cet objectif ou déterminer des exigences spécifiques à chaque classe et en vérifier l'atteinte à la fin de celle-ci. Par ailleurs, dans le programme d'économie familiale, le développement de l'habileté à préparer des mets simples et nutritifs ne se retrouve que dans le module 2 alors que l'on vise de façon globale à développer l'habileté à appliquer le processus de gestion dans tous les modules, ce qui pourrait justifier une épreuve-synthèse à la fin du programme.

Les programmes comprennent des apprentissages de types variés. Il est donc essentiel d'assurer la congruence entre la tâche demandée à l'élève lors de l'épreuve et les caractéristiques de l'apprentissage visé. Ainsi on s'y prendra différemment pour évaluer l'atteinte des objectifs terminaux suivants (programme de biologie humaine de 3e secondaire):

a) «Démontrer la similitude de fonctionnement entre les organes des sens».

b) «Critiquer la composition d'un menu en fonction des besoins qualitatifs de l'adolescent».

Par rapport au premier objectif, il faudrait demander à l'élève de montrer qu'il a compris la similitude de fonctionnement entre les organes des sens. En ce qui a trait au deuxième, on lui demanderait d'évaluer un menu.

Le programme d'études détermine donc les objets d'évaluation, les moments où l'on doit procéder à l'évaluation et parfois la façon de mesurer.

D. Les méthodes d'enseignement

Les méthodes d'enseignement ne devraient pas influencer l'évaluation sommative puisqu'elles ne servent qu'à favoriser l'apprentissage chez l'élève. Ainsi, en géographie 3e secondaire, qu'on s'y prenne d'une façon ou d'une autre pour amener l'élève à être capable de «relier les ressources forestières aux facteurs climat et végétation», cela ne devrait pas influencer la mesure de l'atteinte de cet objectif.

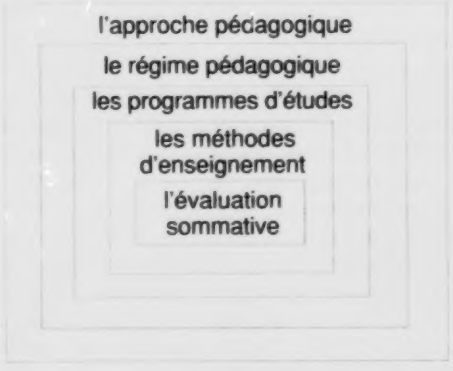
Quelle que soit la méthode d'enseignement utilisée, les objets soumis à l'évaluation sommative ne changent pas. Toutefois, deux sources d'erreur peuvent exister. Premièrement, l'utilisation d'une méthode peut impliquer la poursuite d'objectifs implicites comme l'autonomie, la résolution de problème, l'intégration d'une façon de faire ou encore une conception de l'apprentissage différente de celle qui est sous-jacente au programme. Deuxièmement, la méthode d'enseignement conditionne le vécu de l'élève, le type d'exemples auquel il est confronté, le matériel pédagogique auquel il est habitué, c'est en fait le contexte immédiat dans lequel il baigne et par lequel il développe ses habiletés. D'une part, l'enseignante ou l'enseignant a la responsabilité de faire en sorte que la méthode utilisée corresponde aux objectifs du programme. D'autre part, en évaluation sommative, lors de l'étape mesure, il faut se prémunir contre les particularités des méthodes (objectifs implicites, vécu spécifique). Il n'en demeure pas moins que si cette mesure s'éloigne trop des méthodes utilisées, un élève peut éprouver des difficultés accrues pour faire la démonstration de sa maîtrise. Il est difficile de concevoir une mesure complètement exempte de biais sous ce rapport.

Bref, quels que soient les différents éléments qui peuvent influencer l'évaluation sommative, celle-ci demeure toujours une démarche qui comporte les trois étapes suivantes: prise d'information, jugement et décision. Dans tous les cas, l'enseignante ou

l'enseignant doit s'imprégner des finalités et des objectifs importants des programmes et s'en servir comme points de repère pour l'évaluation sommative.

Il faut cependant être conscient du contexte dans lequel s'exerce l'évaluation sommative: les approches pédagogiques, le régime pédagogique, les programmes d'études et les méthodes d'enseignement .

Le shéma suivant illustre les interrelations entre ces éléments fondamentaux. L'approche pédagogique peut influencer le régime pédagogique parce qu'elle implique une conception particulière de l'homme et de sa place dans la société. Par ailleurs, si la société cherche à assurer une certaine uniformité, comme c'est le cas actuellement, elle établit un régime pédagogique et permet à plusieurs approches pédagogiques de s'y insérer à la suite de modifications et d'adaptations plus ou moins grandes.



2.2 Pour qui l'évaluation sommative est-elle importante?

A. Les élèves

De tous, c'est l'élève qui est le plus concerné par l'évaluation sommative. L'évaluation sommative a un impact sur la perception que l'élève a de lui-même en tant qu'apprenant (concept de soi). Cet impact s'ajoute à celui qu'a eu, pour le même élève, l'évaluation formative. Est-il quelqu'un qui réussit à apprendre? Est-il quelqu'un qui est capable d'apprendre? De plus, l'évaluation sommative peut affecter la motivation de l'élève, facteur important dans son développement.

Les décisions prises en évaluation sommative ont des conséquences cruciales à long terme sur la carrière scolaire et la carrière professionnelle de l'élève. En effet, selon les résultats d'évaluation sommative, on dirigera l'élève vers un cheminement particulier (année de transition, encadrement particulier pour les doués ou pour les retardaires,...), on l'acceptera dans tel cours qu'il ou qu'elle a choisi, on lui donnera son diplôme, on l'admettra au collège ou à l'université,... Puis selon sa carrière scolaire, il ou elle pourra ou non choisir et obtenir un emploi qui correspond à ses intérêts et à ses aptitudes.

À long terme, l'évaluation sommative est donc très importante pour l'élève. Il est essentiel que le jugement exprimé à propos de la maîtrise par l'élève des apprentissages décrits dans le programme, représente la réalité avec le plus de justesse et de pertinence possible.

B. Les parents

C'est surtout l'évaluation formative qui fournit au parent les informations dont il a besoin pour bien assumer ses responsabilités face à son enfant. Toutefois le parent a des

décisions à prendre concernant la carrière scolaire de son enfant. Au moment de prendre ces décisions, il tient compte des informations et des jugements provenant de l'évaluation sommative. Il peut en discuter avec l'enseignante ou l'enseignant, les professionnels ou les administrateurs concernés, lors des rencontres du comité multidisciplinaire ou à d'autres occasions. De cette façon, le parent peut avoir des attentes réalistes et guider son enfant dans son orientation et sa progression dans le système scolaire. Il peut aussi vérifier la qualité de l'évaluation à laquelle son enfant est soumis.

C. Les enseignants

Les enseignants bénéficient eux aussi de l'évaluation sommative. En effet, la fin de l'année est le moment privilégié où l'enseignante et l'enseignant peuvent apprécier globalement l'enseignement donné à la lumière du niveau de compétence auquel arrivent leurs élèves. C'est ainsi qu'il est posssible d'évaluer d'année en année entre autres, la pertinence du contenu, l'organisation de l'enseignement, l'efficience de la séquence des apprentissages appliquée en classe, la précision et le validité des moyens de mesure utilisés. L'évaluation sommative permet enfin de confirmer ou d'infirmer les jugements portés tout au cours de l'année sur les apprentissages des élèves. Elle permet aussi à l'enseignante et à l'enseignant de comparer leurs résultats à ceux obtenus par d'autres. Cela est d'ailleurs favorisé par l'utilisation d'instruments de mesure uniformes.

D. Les professionnels

Souvent les conseillères ou les conseillers pédagogiques sont appelés à participer à l'élaboration d'épreuves uniformes. L'analyse des résultats de ces épreuves permet de vérifier l'application des programmes, ce dont ils tiendront compte par la suite dans leurs interventions.

De plus, les résultats d'évaluation sommative constituent une des informations importantes que les conseillères et conseillers en orientation utilisent lors de rencontres avec les élèves.

E. Les administrateurs

Les résultats d'évaluation sommative sont utiles aux autorités administratives. En étant informées du niveau de maîtrise des apprentissages des élèves et en tenant compte d'un ensemble d'informations complémentaires, celles-ci peuvent effectuer les classements, organiser l'année scolaire, allouer les ressources selon les besoins. De plus, les personnels de direction (d'écoles, de services éducatifs, de commissions scolaires,...) ont le devoir de répondre à la société du fonctionnement de l'institution, de leur façon d'utiliser les fonds publics et de la qualité des services dispensés. Ils doivent entre autres s'assurer que les élèves soient soumis à des formes valables d'évaluation. L'évaluation sommative offre des points de repère stables d'une année à l'autre et reconnus valables pour juger de la compétence de l'élève. De ce fait, elle est essentielle à une gestion efficace et sûre du système scolaire.

L'évaluation sommative est donc utile aux différents agents éducatifs, que ce soit les élèves, les parents, les enseignants ou les administrateurs. Les conséquences de ce type d'évaluation touchent d'abord l'élève: le cours de sa vie en est directement influencé. Pour les autres agents éducatifs l'évaluation sommative a aussi une importance certaine.

2.3 Comment partager les responsabilités?

Le ministère de l'Éducation émet des directives concernant le partage des responsabilités. Il appartient d'autre part à la commission scolaire de déterminer les rôles de

chacun de ses membres et de distribuer les tâches au moment de la planification de l'évaluation sommative. C'est ainsi que les enseignants, les professionnels et les administrateurs pourront être impliqués dans les situations d'évaluation, chacun selon sa compétence.

A. Le ministère de l'Éducation

Le ministère de l'Éducation a la responsabilité de la sanction des études, il y contribue de façon plus importante en 4e et 5e secondaire. Comme c'est lui qui décerne les diplômes, il doit donc assurer la qualité de l'évaluation. Pour ce faire, il consacre des ressources pour mettre au point des épreuves de qualité. Par la suite, il peut fournir à tous les organismes en plus des résultats locaux, les résultats provinciaux, ce qui donne alors un point de repère utile aux commissions scolaires, aux professionnels et aux enseignants.

B. La commission scolaire

La commission scolaire a la responsabilité d'établir les règles de passage d'une classe à l'autre incluant le passage du primaire au secondaire. C'est elle aussi qui doit gérer l'attribution des crédits pour les premières années du secondaire.

La commission scolaire doit aussi assurer la qualité de l'évaluation sommative dont elle est responsable. C'est dans cet esprit qu'elle planifie la démarche d'évaluation, partage les responsabilités, alloue des ressources et implique les enseignants et les professionnels.

C. Le personnel enseignant et professionnel

Qu'il s'agisse de la sanction des études, du placement dans un cheminement particulier ou du passage à une autre classe, on sollicite toujours la recommandation de l'enseignante ou de l'enseignant. Parce qu'elle connaît le vécu scolaire de l'élève pour l'année, c'est cette personne qui peut le mieux juger de la compétence de l'élève. Il est possible alors de tenir compte entre autres des progrès de l'élève, du travail effectué, des difficultés momentanées ou durables en plus de ses résultats scolaires, en rapport avec la décision à prendre.

Par contre, cette subjectivité est une arme à double tranchant. Elle constitue une richesse puisqu'elle permet par exemple d'identifier une note d'examen non représentative de l'apprentissage d'un élève. En fait, elle est susceptible d'amener à une connaissance assez juste de l'élève, mais comme elle suppose une implication affective de l'enseignante ou de l'enseignant, elle ouvre la voie à l'influence de perceptions erronées, de stéréotypes.

C'est ainsi dire qu'on pourrait être porté à juger qu'un élève peu travailleur, dissipé et indiscipliné ne maîtrise pas les apprentissages visés (Weiss, 1984). La prudence s'impose donc.

Pour pallier cette difficulté, certains enseignants ou enseignantes ont tendance à se regrouper au niveau de l'école. Ils préparent des examens communs qui leur apportent une base plus objective leur permettant de porter de meilleurs jugements. Ils se réunissent en comité pour discuter de cas, comité multidisciplinaire pouvant regrouper les parents, les professionnels et la directrice ou le directeur concerné.

Les conseillers pédagogiques et les spécialistes de la mesure en éducation sont sussceptibles de participer à ces comités. Ils fournissent surtout leur expertise lors de l'élaboration d'examens uniformes; cette expertise concerne généralement le choix du contenu des épreuves et la méthode d'élaboration de celles-ci.

À plus long terme, une commission scolaire favorise l'amélioration de la pratique de

l'évaluation sommative par l'affectation de ressources au perfectionnement en évaluation de son personnel enseignant.

Bref, les divers agents impliqués en évaluation sommative des apprentissages, en mettant en commun leurs différentes ressources, peuvent en arriver à porter un jugement juste sur la maîtrise des apprentissages par l'élève et à prendre les décisions pertinentes. Il est donc important de clarifier les responsabilités de chacun.

2.4 Résumé

Certains éléments du contexte dans lequel s'effectue l'évaluation sommative sont susceptibles de l'influencer dans une ou plusieurs de ses étapes. Le régime pédagogique et les programmes d'études déterminent directement l'application de l'évaluation sommative. L'approche pédagogique pour sa part peut l'influencer mais dans les limites imposées par le régime pédagogique. C'est à l'étape mesure qu'on est le plus confronté aux particularités des méthodes d'enseignement, il s'agit alors d'utiliser des instruments qui favorisent la démonstration de la maîtrise chez tous les élèves concernés.

L'évaluation sommative est importante pour tous les agents de l'éducation mais à des degrés différents et sous des aspects variés. C'est de toute évidence chez l'élève que l'impact est le plus personnel et le plus direct. Les parents, les enseignants, les professionnels et les administrateurs pour leur part utilisent les résultats d'évaluation sommative pour mieux guider l'élève, faire un retour sur leur action et planifier leurs prochaines interventions à long terme.

Les responsabilités sont partagées entre les divers agents de l'éducation selon les dispositions déterminées par la commission scolaire et le ministre. Ces reponsabilités varient selon qu'il s'agit de l'ordre d'enseignement primaire ou secondaire, le ministère de l'Éducation étant plus impliqué au second cycle du secondaire.

2.5 Références

ASSEMBLÉE NATIONALE DU QUÉBEC. *Loi sur l'enseignement primaire et secondaire public*, Projet de loi 3 (1984, chapitre 39), Éditeur officiel du Québec, 1984, 137 p.

BEAUDOIN, Jean-Marc. «La décision du juge Lebrun. La C.S.T.R. est satisfaite mais ne crie pas victoire», in *Le Nouvelliste*, 10 mars 1984, p. 32

BERTRAND, Yves, et Paul, VALOIS. *Les Options en éducation*, 2e éd. revue et corrigée, Gouvernement du Québec, Ministère de l'Éducation, Direction de la recherche, Québec, 1982, 191 p.

BIRZEA, César. *La pédagogie du succès*, Paris, Presses universitaires de France, 1982, 152 p.

BLOOM, Benjamin S, J. Thomas, HASTINGS, et Georges F., MADAUS. *Handbook on Formative and summative Evaluation of Student Learning*, New-York, McGraw-Hill, 1971, 923 p.

BLOOM, Benjamin S, Georges F., MADAUS, et J. Thomas, HASTINGS. *Evaluation to Improve Learning*, Mc Graw-Hill Book Company, New York, 1981, 356 p.

CARDINET,Jean. *Compte rendu du groupe de travail sur l'articulation de l'évaluation sommative et formative*, Actes des rencontres, Évaluation formative et formation des enseignants, Namur, 1983, p. 102 à 105.

CARDINET, Jean. *Des instruments d'évaluation pour chaque fonction*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1983, 81 p.

CARDINET, Jean. *Réflexions d'enseignants sur l'évaluation des élèves*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1984, 25 p.

CARDINET, Jean. *Pour apprécier le travail des élèves,* Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1984, 75 p.

GARIÉPY, Wilfrid. *Pour la supervision de l'enseignement, l'évaluation de l'apprentissage*, Montréal, Association des institutions d'enseignement secondaire, 1973, 60 p.

GOUVERNEMENT DU QUÉBEC, Conseil supérieur de l'Éducation. *L'Évaluation. Situation actuelle et voies de développement*, Rapport 1982-1983 sur l'état et les besoins de l'éducation, Québec, 1983, 132 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation. *Politique générale d'évaluation pédagogique*, Québec, 1981, 23 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Règlement concernant le régime pédagogique du primaire et l'éducation préscolaire*, 16-0063-01, Québec, 1981, 54 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Règlement concernant le régime pédagogique du secondaire*, 16-0062-01, Québec, 1981, 54 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Programme d'études. Géographie du Québec et du Canada*, 3e secondaire, Formation générale et professionnelle, Québec, avril 1982, 67 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Programme d'édudes secondaires. Biologie humaine*, Québec, août 1982, 52 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation du Québec. *L'école québécoise: Énoncé de politique et plan d'action*, Québec, 1er trimestre 1979, 163 p.

LÉVESQUE, Marcienne. Relation entre la pédagogie pour la maîtrise de l'apprentissage (P.M.A.) et le concept de soi scolaire chez des élèves de première secondaire, Thèse de doctorat inédite, présentée à l'université Laval, 1982, 304 p.

NITKO, Anthony J. *Educational Tests and Measurement. An Introduction,* Toronto, Harcourt brace Jovanovich, 1983, 661 p.

PARÉ, André. *Créativité et pédagogie ouverte.* Volume III: *Organisation de la classe et intervention pédagogique*, Editions N.H.P., 1977, 343 p.

SHEPARD, Loretta A. «Setting Standards and Living with Them», *Florida journal of Educational Research*, 1976.

THURLER, Monica. *Réflexions sur l'évaluation de l'élève,* Rapport d'atelier I, Groupe de travail 2, Genève, Directeurs du projet SIPRI, février 1983, 48 p.

WEISS, Jacques. *La subjectivité blanchie?* Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1984, 12 p.

Comment planifier l'évaluation sommative?

Préambule 23
3.1 Planifier l'étape décision 25
A. Identifier le type de décision à prendre 25
B. Faire l'inventaire des possibilités . . 26
C. Établir les règles à suivre 27
3.2 Planifier l'étape jugement 27
A. Comment juger de la compétence d'un élève en regard d'un programme d'études? 28
B. Comment juger des autres éléments pertinents à l'orientation de l'élève dans sa carrière scolaire? 32
C. Comment traiter les cas limites et les cas surprises? 33
D. Déterminer les moments où l'on portera les jugements 34
3.3 Planifier l'étape mesure 34
A. Choisir la façon d'obtenir les informations nécessaires aux jugements à porter 36
B. Choisir les façons d'interpréter les informations et d'établir le seuil de réussite s'il y a lieu 39
C. Déterminer les moments où l'on procédera à la mesure 44
3.4 Choisir une façon de conserver les données 45
3.5 Résumé 47
3.6 Références 48

Chapitre troisième

Dans ce chapitre, nous verrons que:

1. Il est très important de planifier le déroulement de l'évaluation sommative.
2. C'est le type de décision à prendre qui oriente la démarche d'évaluation sommative tout entière.
3. Les règles de décision spécifient ce dont on doit tenir compte selon le type de décision à prendre, les jugements dont on aura besoin et les personnes qui prendront les décisions.
4. On peut faire deux types d'erreurs en prenant une décision: ne pas prendre telle décision qu'on aurait dû prendre et prendre telle décision qu'on n'aurait pas dû prendre.
5. On peut regrouper les décisions d'après deux catégories selon l'intention qui les guide: reconnaître les acquis obligatoires ou prédire le succès.
6. Connaître les ressources du milieu permet d'établir des règles de décision pertinentes et de veiller à obtenir les jugements nécessaires.
7. Les jugements s'appuient sur les résultats de diverses prises d'informations (mesures, observations,...).
8. Pour juger si un élève a atteint un niveau suffisant, on compare le niveau d'apprentissage global de l'élève à celui fixé par le programme.
9. Il existe plusieurs techniques pour spécifier les exigences d'un programme.
10. Il y a plusieurs précautions à prendre pour éviter les erreurs dans les jugements.
11. Il est préférable de juger de la compétence d'un élève le plus près possible de la fin de la période prévue pour les apprentissages visés.

12. On doit planifier l'étape jugement en fonction de la décision à prendre.
13. Les informations sur lesquelles sont basés les jugements doivent être de qualité (valides et fidèles).
14. Dans la planification des prises d'informations, on précise les informations nécessaires en fonction du type de décision à prendre et des jugements à porter.
15. Il existe plusieurs types d'épreuves auxquelles on peut avoir recours pour obtenir l'information nécessaire aux jugements à porter.
16. Le seuil de réussite est le niveau de qualité à partir duquel on peut interpréter une performance comme reflétant la compétence.
17. On doit tenir compte du fait qu'il est impossible de savoir si le résultat d'un élève représente parfaitement son habileté.
18. Il faut prévoir conserver les informations prises et les jugements temporaires posés.
19. Lorsqu'on prévoit l'échéancier, il est prudent de se garder une marge de manoeuvre pour bien préparer chaque étape.

Préambule

De façon globale, planifier c'est déterminer des buts à atteindre, des moyens, un scénario de réalisation et de partage des tâches.

En d'autres mots, quand on planifie, on prévoit et on relie d'une certaine manière des moyens à un but que l'on vise.

Il en va de même quand on parle de la planification de l'évaluation sommative:

- on considère les décisions en jeu et ce dont on a besoin pour les prendre,
- on détermine les moyens d'obtenir ce dont on a besoin,
- on établit un calendrier d'opérations ainsi que le partage des tâches.

C'est le type de décision à prendre qui oriente la démarche tout entière, le choix des informations dont on tiendra compte et les jugements à poser. Nous avons vu que les décisions possibles se rapportent à la sanction des études d'un élève, à la reconnaissance de ses acquis expérientiels, à son passage à la classe supérieure et à son orientation.

D'une part, si l'on veut être juste dans la sanction des études, il faut établir des règles claires, assurer la cohérence entre les apprentissages qui sont à évaluer et ceux déterminés par les programmes: c'est se donner une définition stable de la réussite. D'autre part, s'il s'agit de décider de la promotion d'un élève ou de son placement dans un groupe, il faut être en mesure de prévoir la situation qui lui sera le plus favorable: c'est se donner une définition stable de la prévision de la réussite.

Toutes ces étapes de la planification supposent beaucoup de travail. Elles devraient être terminées avant le début de l'enseignement. Comment en tant qu'enseignant pourra-t-on orienter son action si les règles du jeu ne sont pas définies à l'avance? Comment pourra-t-on être plus certain de la justesse des décisions concernant l'élève? Les décisions impliquées dans l'évaluation sommative ont des conséquences importantes. Même s'il arrive qu'on modifie la décision prise, on ne peut penser annuler complètement les effets que la première décision a eu le temps de provoquer. Il vaut donc mieux prendre le temps de bien planifier pour favoriser le plus possible la qualité des décisions.

D'autres facteurs incitent aussi à planifier cette démarche:

— l'obligation d'informer les parents et les élèves sur la façon d'évaluer et sur les conditions de passage à la classe supérieure;

— les recommandations qu'il faut donner au cours de la deuxième moitié de l'année scolaire concernant le passage à une classe supérieure et les façons de regrouper les élèves;

— l'envergure de l'organisation à mettre sur pied (élaboration d'instruments de mesure communs, formation de comités,...);

— le temps à investir pour arriver à des consensus, élaborer les examens, apporter les améliorations nécessaires, veiller à l'organisation matérielle...

— la diversité des agents impliqués dans la démarche.

De façon plus précise, lors de la planification de l'évaluation sommative on doit établir:

— les décisions à prendre;

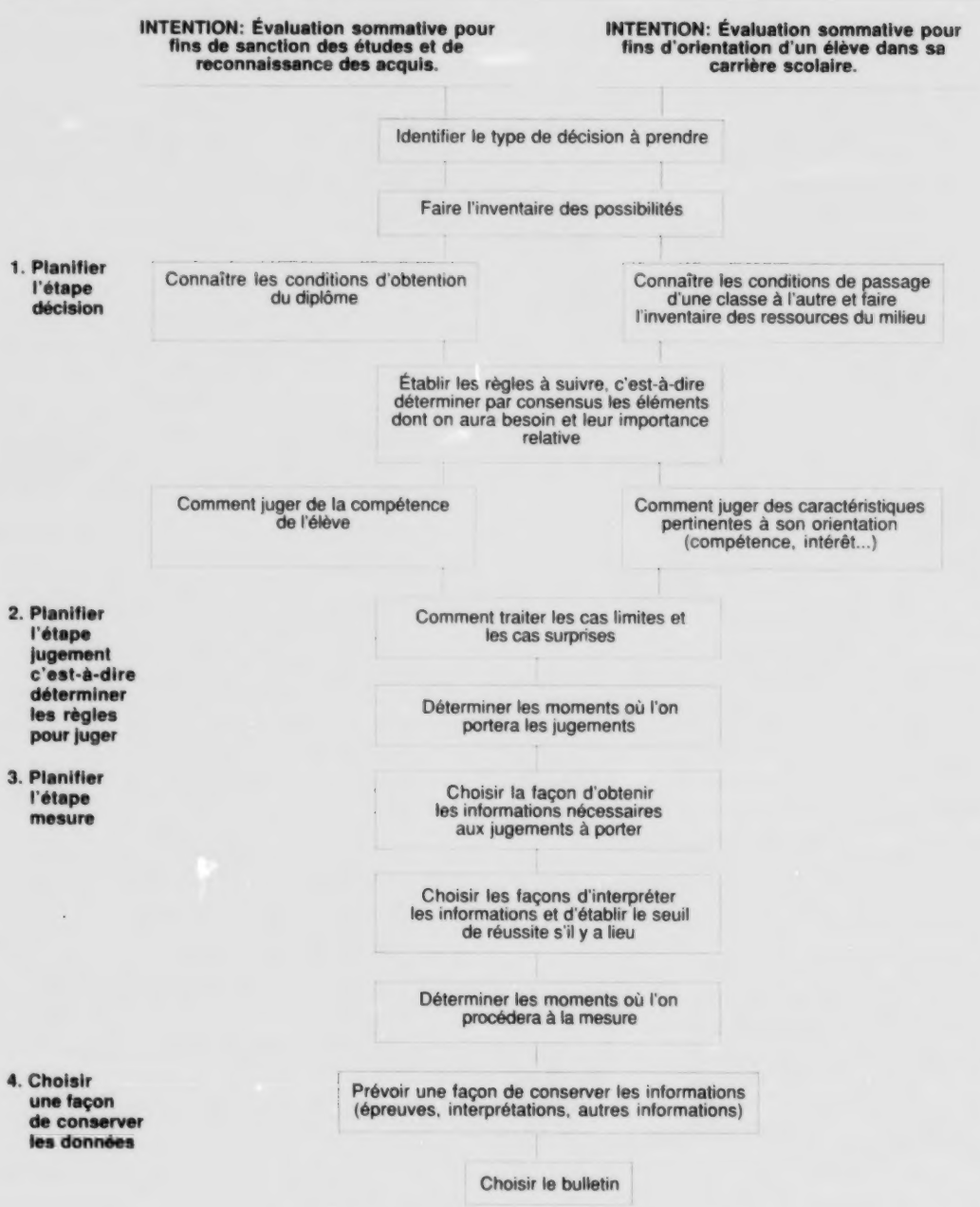
— les règles pour juger;

— les façons d'obtenir l'information pertinente;

— les façons de conserver les données.

Le tableau 2 présente les différentes étapes de la planification de l'évaluation sommative.

Tableau 2
La planification de l'évaluation sommative



3.1 Planifier l'étape décision

Pour établir les règles de décision, on se demande:

- de quoi on doit tenir compte dans le type de décision à prendre,
- de quels jugements on a besoin,
- qui prendra la décision.

Quel que soit la décision à prendre, deux types d'erreurs sont possibles. Par exemple, dans le tableau 3, la décision en jeu concerne l'octroi de crédits. L'erreur survient quand la décision prise ne correspond pas à l'état réel de l'élève. On fera une erreur de type I si on n'accorde pas les crédits à un élève qui a réalisé les apprentissages et qui aurait droit d'obtenir ces crédits. On fait aussi une erreur, de type II cette fois, quand on accorde les crédits à un élève qui n'a pas réalisé les apprentissages. Par exemple, on peut décider de refuser de placer un élève qui en a besoin dans une classe de transition ou encore, décider d'y placer un élève qui n'en a pas besoin. C'est pourquoi l'étape de planification est si importante. Mieux on pèse les conséquences de ces deux types d'erreurs, plus on fait preuve de prudence lors de la prise de décision. On doit s'interroger sur les coûts et bénéfices attachés à chaque type d'erreur. On tente d'identifier les conséquences pour l'élève (son présent et son futur) de même que pour la société (dévaluation du diplôme, durée des études...). Par exemple, dans le cas d'un élève qui aurait atteint presque le seuil de réussite la question suivante se pose: les conséquences sont-elles plus graves si on le fait doubler ou si on le fait passer à la classe supérieure?

La planification des règles de décision est importante car elle donne le temps de faire des choix plus réfléchis que ceux que l'on pourrait faire dans le feu de l'action. Elle peut ainsi aider à éviter ou du moins à diminuer des erreurs coûteuses tant sur le plan individuel que social. Sans planification, la justice en évaluation sommative devient certes plus aléatoire.

A. Identifier le type de décision à prendre

La première étape de la planification consiste à identifier le type de décision à prendre ainsi que les alternatives possibles.

Il existe quatre types de décisions:

a) sanctionner les études: décerner les diplômes et octroyer les crédits;
b) reconnaître les acquis expérientiels;
c) décider du passage à la classe supérieure;

Tableau 3

Types d'erreurs possibles

Décision prise	**État réel de l'élève**	
	L'élève a réalisé les apprentissages	L'élève n'a pas réalisé les apprentissages
Je n'accorde pas les crédits	Mauvaise décision: Erreur de type I (faux-négatif)	Bonne décision
J'accorde les crédits	Bonne décision	Mauvaise décision: Erreur de type II (faux-positif)

d) orienter l'élève dans sa carrière scolaire: cheminement particulier, choix d'options, etc.

On peut regrouper ces types de décisions en deux catégories selon l'intention qui nous guide:

- nous assurer que l'élève est compétent (sanction et équivalences);
- nous assurer que l'élève aura les conditions optimales de réussite(promotion et orientation).

Il est important de distinguer ces deux catégories d'intentions car on se base sur des jugements et des informations différentes selon le type de décision que l'on prend.

Par exemple, quand il s'agit d'accorder le diplôme d'études secondaires, on s'assure que l'élève a fait les apprentissages des programmes d'études jugés importants par consensus en vertu d'une approche critériée et qu'il rencontre les règles déterminées par le ministère de l'Éducation pour la sanction des études. D'autre part, si l'on veut placer l'élève dans la situation qui lui sera le plus propice (doubler, choisir une option...), on tient compte de ses acquis bien sûr, mais aussi de ses intérêts, de sa personnalité, etc. Ainsi De Ketele (1983) a rapporté que les facteurs cognitifs ne comptent au maximum que pour 25% dans l'explication de la performance dans les études supérieures. Si on y ajoute l'intérêt pour les études, la prédiction passe à 45%. Si en plus, l'orientation de l'individu est le fruit de sa décision personnelle et qu'il est capable de se fixer des objectifs de travail et de les ajuster selon les événements, la prédiction de la performance monte à 80%. Par facteurs cognitifs, on entend les aptitudes intellectuelles, la réussite académique antérieure, les capacités cognitives de base et les styles cognitifs.

B. Faire l'inventaire des possibilités

Une fois l'intention précisée, on fait l'inventaire des possibilités. On se demande: Quels sont les choix possibles?

Le premier type de décision concerne la reconnaissance de la compétence de l'élève . Les choix sont:

— d'accorder ou non les crédits à l'élève;
— de reconnaître ou non ses acquis expérientiels.

On devra donc faire l'inventaire des conditions d'obtention du diplôme en se référant aux documents officiels.

Le deuxième type de décision se rapporte à la carrière scolaire de l'élève. Les choix sont de:

— placer l'élève dans un regroupement particulier (programme pour élèves doués, cours de récupération avec plus de temps dans la matière, année de transition, enseignement individualisé, classe à degrés multiples, aide d'une ou d'un orthopédagogue, programme pour décrocheurs, école alternative, intégration de matières, cours optionnel, périodes d'études organisées à l'école, programme particulier pour un groupe d'élèves pour qui l'enseignement d'une matière a été perturbé au cours de l'année tout en les faisant monter...);
— faire sauter une ou plusieurs classes;
— faire doubler une classe ou recommencer un programme;
— faire suivre des cours d'été;
— faire passer l'élève à la classe supérieure (avec appui, sans appui).

On devra donc faire l'inventaire des ressources du milieu et des conditions de passage à la classe supérieure. Ceci nous permet d'établir des règles de décision pertinentes et de veiller à obtenir des jugements utiles à ces prises de décision. Par

exemple, s'il n'y a pas de cours d'été offerts dans la région, on ne peut fixer une règle qui serait de faire monter des élèves un peu faibles à la condition qu'ils réussissent des cours d'été, à moins bien sûr, de décider de mettre en place de tels cours.

C. Établir les règles à suivre

Une fois que l'on a identifié la décision à prendre, on choisit par consensus les éléments dont on aura besoin pour prendre cette décision, on établit les règles de décision pour les cas limites.

Par exemple, on se demande de quels **éléments** on tiendra compte lors du passage de la troisième à la quatrième année primaire. Ce pourrait être de:

- l'état des apprentissages de l'élève;
- son degré de maturité intellectuelle;
- son degré de maturité psychosociale.

On pourrait alors établir les règles qui suivent:

- si la situation de l'élève est satisfaisante pour ces trois éléments, il passe en quatrième année;
- si la situation de l'élève n'est pas satisfaisante pour ces trois éléments, il reste en troisième année;
- si l'état de ses apprentissages est légèrement insatisfaisant mais que l'élève a un degré de maturité adéquat, il passe en quatrième avec quelques mesures d'appui pour ses faiblesses spécifiques;
- si l'état de ses apprentissages est très insatisfaisant, il reste en troisième année si on prévoit que, d'après ses caractéristiques psychologiques, il réagira mal s'il est placé dans une classe de transition.

Il s'agit donc de prévoir comment on prendra les différentes décisions en tenant compte des éléments que l'on considère pertinents et de la combinaison de ces éléments chez le même individu qui fera que telle décision aura des conséquences favorables pour cet élève et pas nécessairement pour tel autre.

L'étendue et la diversité des éléments dont on voudra tenir compte varient selon la décision à prendre et le contexte dans lequel cette décision se situe. Pour décider du passage à une classe supérieure au primaire ou à un cycle supérieur au secondaire, surtout quand l'enseignement se donne dans une autre école, on utilise un ensemble d'éléments différent de celui auquel on se réfère lorsqu'il s'agit d'accorder des crédits pour la réussite à un cours précis.

Les jugements que l'on porte sur la situation de l'élève par rapport à chaque élément estimé important lors de la planification des décisions s'appuient sur les résultats des diverses prises d'informations effectuées. Ces résultats peuvent provenir de mesures proprement dites comme les interprétations faites à partir des performances à des épreuves-synthèse et de certaines observations qui permettent de mieux juger de ces résultats ou qui apportent d'autres informations. Par exemple, ces observations pourraient porter sur le comportement psychosocial de l'enfant ou encore sur certains événements précis qui ont pu le perturber (maladie, déménagement, décès dans la famille, etc.).

3.2 Planifier l'étape jugement

Après avoir arrêté la façon dont on prendra les décisions, on détermine comment on fera pour juger de la situation de l'élève par rapport aux divers éléments retenus, c'est la deuxième des quatre étapes de la planification de l'évaluation sommative. Cette façon de faire doit être en rapport avec la décision en cause. On spécifie comment juger de la compétence de l'élève ou des autres caracté-

ristiques pertinentes à son orientation dans sa carrière scolaire. On prévoit également comment traiter les cas limites et les cas surprises. On détermine aussi les moments où l'on portera les jugements.

Dans la plupart des décisions en évaluation sommative, on accorde une grande importance au jugement sur la compétence par rapport aux apprentissages importants au programme. À première vue, cette définition des apprentissages importants donne l'impression qu'il s'agit de points de repère absolus, d'éléments stables de contenus académiques. Mais il ne faudrait pas croire qu'il s'agit d'un absolu où l'élève est seul devant la matière. Toute définition de ces apprentissages à un niveau ou à un âge donné passe par une étude comparative qui implique un ensemble d'élèves, c'est-à-dire par une étude de normalisation. Autrement dit, en évaluation sommative, on a besoin de comparer ce dont l'élève est capable avec ce qui a été reconnu, après expérimentation, comme étant possible avec des élèves du même niveau dans le cadre du même programme.

Dans les faits, il s'avère que les évaluatrices et les évaluateurs utilisent des critères très variés lorsqu'ils portent des jugements en évaluation sommative:

— le degré de maîtrise absolue du domaine ou des objectifs par rapport à la performance du groupe d'élèves;
— la position de l'élève dans le groupe;
— la qualité des expériences d'apprentissages;
— les attitudes des élèves;
— la qualité des examens;
— le cheminement du groupe d'élèves par rapport aux autres groupes;
— le regroupement des élèves dans des rythmes, des programmes spéciaux, etc.;
— l'importance relative accordée aux critères explicites;
— les valeurs (religieuses ou morales) de l'institution ou du personnel enseignant;
— les règlements disciplinaires;
— les politiques provinciales, régionales, locales (âge de passage du primaire au secondaire...);
— les contraintes humaines, financières, physiques (Morissette: 1984 p. 39 à 41).

En plus d'être peu uniformes, ces critères ne seraient pas toujours explicités, ni connus des élèves. Il est donc fort important de planifier l'étape jugement afin d'assurer la reconnaissance des critères choisis et leur acceptation.

A. Comment juger de la compétence d'un élève en regard d'un programme d'études?

Juger de la compétence d'un élève suppose d'une part, qu'on fasse une description de cette compétence et d'autre part, qu'on tienne compte de toutes les informations pertinentes. Au moment de la planification, on doit choisir et faire connaître la définition de la compétence qui servira de point de référence selon une approche critériée et indiquer les autres facteurs dont on tiendra compte.

Lorsqu'il s'agit de sanction des études ou de reconnaissance des acquis expérientiels, l'approche critériée nous amène à vérifier si l'élève possède l'ensemble des habiletés et des connaissances prescrites dans les programmes, alors qu'une approche normative nous conduirait à juger les élèves les uns par rapport aux autres. Les programmes d'études comprennent des objectifs d'ordre affectif (attitudes, intérêts, valeurs) mais la pertinence d'en tenir compte en évaluation sommative n'est pas encore établie. Enfin, les opinions sont très partagées et la question

n'a pas été suffisamment formulée pour que l'on puisse conseiller une façon de faire.*

Pour juger si l'élève a réalisé les apprentissages prévus dans un programme, il faut comparer l'ensemble des apprentissages réalisés par l'élève aux exigences établies par consensus.

Ce travail de clarification du programme peut être fait en partie par le ministère de l'Éducation, par la commission scolaire et par l'enseignante ou l'enseignant de préférence en équipe. L'important est de préciser des exigences réalistes par rapport au programme en question dans le cadre de la décision à prendre. Pour ce faire on tient compte du contexte. Ainsi dans le cas des programmes élaborés par cycle, on doit commencer par répartir le contenu par classe en tenant compte du milieu. Dans une région où l'enseignement de l'anglais n'est pas très poussé au primaire, les exigences à la fin de la 1ère secondaire doivent être réalistes avec le niveau d'entrée des élèves. S'il s'agit d'un programme comportant des modules au choix (Initiation à la technologie), il faut évidemment que l'évaluation porte sur les modules qui seront enseignés. On devra aussi tenir compte du fait qu'un apprentissage se continue dans un autre programme, est préalable à un programme subséquent ou est final. L'établissement des exigences pour la sanction des études suppose qu'on ait précisé la définition de ce qui est essentiel selon les finalités de l'éducation.

Chose certaine, plus l'établissement des exigences se fait de façon ouverte, moins il y a de risques qu'elles soient rejetées par la suite. C'est-à-dire que si les enseignants ou les conseillers pédagogiques par exemple y ont participé, ils en auront une bonne compréhension et ils devraient normalement les accepter.

De plus, on devrait réviser annuellement ces exigences pour tenir compte de l'évolution du milieu. À l'étape de la planification, il demeure que la règle qu'on établit n'a pas été confrontée à la réalité, elle est l'image des prévisions que l'on fait. Il est prudent d'en prévoir la révision lors de l'étape où l'on jugera vraiment chaque cas d'élève. On analysera alors les conséquences de cette règle. On vérifiera, par exemple, si la règle établie a pour conséquence de faire échouer une grande proportion des élèves. Il est important que les exigences sur lesquelles on basera ses jugements soient:

— explicites: réellement exprimées, formulées;

— objectivables: plusieurs personnes peuvent arriver à des résultats semblables en employant ces exigences;

— communicables: qui peuvent être facilement comprises.

A-1 Quelques façons de décrire la compétence

Le jugement de la compétence de l'élève nécessite donc le recours à des critères. Plus ces critères sont explicites, plus le jugement a des chances d'être juste et conforme à la réalité. Ce qu'il s'agit de faire alors c'est de déterminer les apprentissages importants d'un programme, les exigences sur lesquelles portera le jugement. Il est entendu qu'une analyse du programme a été faite précédemment lors de la planification de l'enseignement. Les façons de faire que l'on présente ci-dessous sont habituellement employées pour préparer une épreuve. Nous pouvons toutefois les utiliser, en les adaptant, pour choisir les apprentissages sur lesquels on voudra faire porter les jugements. N'oublions pas que les choix qui y seront faits doivent toujours s'appuyer sur la décision à prendre.

* Voir le document *Vers un bulletin scolaire renouvelé*, Gouvernement du Québec. Document d'information 16-7513.

Il existe plusieurs méthodes pour préciser les exigences d'un programme:

a) la définition du domaine pour fins d'évaluation;

b) L'élaboration d'un tableau de spécifications;

c) la détermination de l'importance relative des objectifs du programme.

Ces trois méthodes s'appliquent autant dans une approche critériée que dans une approche normative. Selon l'approche critériée on cherche avant tout à bien représenter le programme pour bien identifier la compétence de l'élève. Selon l'approche normative, la priorité est accordée à la différenciation entre les individus, même si on a le souci de bien représenter le programme. Par la suite, le choix ou l'élaboration des instruments de mesure diffèrent de ceux faits selon une approche critériée. Ainsi, on élimine les questions réussies ou échouées par une grande proportion d'élèves, ce qui n'est pas le cas dans une approche critériée.

a) La définition du domaine

La définition du domaine est une méthode qui permet de déterminer, pour fins d'évaluation, les apprentissages importants d'un programme. Des trois méthodes présentées, c'est la plus complète parce qu'elle tient compte à la fois des orientations du programme d'études, des habiletés et des contenus. Son utilisation favorise donc la validité des jugements qu'on portera sur la compétence de l'élève.

Ici la définition du domaine se réalise en 3 étapes principales*:

1. Énumérer les principes à respecter dans l'évaluation. Ils découlent des orientations ou des priorités du programme déjà analysées lors de la planification de l'enseignement.

2. Identifier et décrire les éléments du programme sur lesquels on compte juger de la compétence des élèves.

3. Regrouper ces éléments essentiels retenus en un tableau ou un schéma qui présente les dimensions et indiquer l'importance que l'on devrait leur accorder lors des jugements à porter pour établir la compétence de l'élève.

* *Guide pour la construction d'instruments de mesure dans le cadre de la Banque d'instruments de mesure*, Chapitres II – III, Documents de travail B.I.M., septembre 1984, 16-7300-02, pp. 2-3.

b) L'élaboration d'un tableau de spécifications

Pour préciser les exigences d'un programme, on peut aussi élaborer un tableau de spécifications car cette méthode n'est pas restreinte à une approche particulière (critériée ou normative), ni à un type d'évaluation spécifique (formative ou sommative). Son utilisation se justifie par le fait que c'est une façon de préciser les apprentissages du programme sur lesquels on veut faire porter les jugements, selon la décision à prendre. De plus, cette méthode est assez connue dans le milieu scolaire. On choisit d'abord parmi les apprentissages prévus au programme ceux sur lesquels on basera l'évaluation sommative. Pour ce faire, on procède à l'analyse de l'ensemble ou d'une partie des apprentissages prévus au programme. On identifie les contenus d'ordre cognitif, affectif ou moteur reliés à ce programme, on les classe par rapport à une taxonomie, puis on indique leur importance relative.

L'exemple ci-dessous est le résultat de l'analyse d'une partie du programme de français de 1ère secondaire, la production écrite. Pour obtenir un tableau d'ensemble de ce programme, il faudrait y combiner les tableaux établis à partir des choix faits au sujet d'autres habiletés aux programmes: production orale, compréhension écrite et orale. Une plus grande importance serait alors accordée à la situation de communication. Il faut noter que de plus en plus, dans le cadre de l'évaluation sommative, la définition du domaine remplace le tableau de spécifications parce qu'elle est le résultat d'une démarche plus complète.

c) La détermination de l'importance relative des objectifs du programme

Une troisième façon de décrire la compétence consiste à élaborer la liste des objectifs importants du programme sur lesquels on veut baser ses jugements et à les pondérer. Il n'est pas nécessaire de mesurer un à un les objectifs qui représentent différents stades d'apprentissage intermédiaires ni les objectifs partiels. On choisit parmi tous les objectifs importants un ensemble représentatif du contenu et de l'importance accordée à chaque objectif selon la décision à prendre. Un niveau d'exigence peut être fixé selon le cas pour chaque objectif. On pourrait aussi en fixer un pour l'ensemble des objectifs par exemple 60% des objectifs importants pour la sanction des études ou encore 75% pour tel programme en vue d'être accepté dans telle option. Partant de cette liste d'objectifs pondérés, on peut procéder de deux façons pour établir les exigences: le profil ou la compensation. Pour étudier ces deux méthodes prenons l'exemple suivant.

Utiliser des profils: supposons que les objectifs 1 et 3 soient condidérés comme essentiels, on pourrait juger que le programme d'études est maîtrisé seulement par les élèves qui ont réalisé ces deux objectifs. Il s'agit évidemment d'objectifs globaux. Dans l'exemple prédédent il n'y aurait que Julie qui réussirait. On pourrait aussi décider qu'un élève qui atteint un des deux objectifs obligatoires et deux autres objectifs a réussi le programme; Julie, Louis et Martine seraient alors considérés comme ayant réussi. D'un autre côté, si un des objectifs obligatoires n'est pas atteint, on peut exiger que les autres objectifs soient très bien maîtrisés.

Exemple d'un tableau de spécifications (Taxonomie de Bloom)
Français 1re secondaire: Production écrite

Contenu	Habiletés						
	Connaissance	Compréhension	Application	Analyse	Synthèse	Évaluation	Total
Intention	—	—	—	—	20%	—	**20%**
Rapport	—	—	—	—	20%	—	**20%**
Lexique	—	—	—	—	6%	—	**6%**
Syntaxe	—	—	30%	—	—	—	**30%**
Orthographe d'usage	—	—	10%	—	—	—	**10%**
Orthographe grammaticale	—	—	14%	—	—	—	**14%**
Total	—	—	**54%**	—	**46%**	—	**100%**

Appliquer le processus de compensation: si on décide d'appliquer le processus de compensation, on n'exige pas que les élèves réussissent les mêmes objectifs. On pourrait par exemple décider que les élèves qui ont réussi 60% des objectifs et plus soient jugés compétents peu importe le profil de ces 60%. Dans l'exemple précédent, Louis, Martine et Manon qui ont obtenu la même moyenne d'objectifs atteints seraient tous trois jugés compétents même si ce ne sont pas les mêmes objectifs qu'ils ont maîtrisés. Pour appliquer le processus de compensation, il ne faut pas que certains des apprentissages au programme soient indispensables à la poursuite de la formation scolaire ou à la pratique du métier visé.

A-2 Tenir compte de toutes les informations pertinentes

L'étape jugement suppose qu'on prenne une distance par rapport à l'interprétation des résultats et qu'on tienne compte de renseignements complémentaires à ceux fournis par la mesure. Ces derniers peuvent concerner:

- la validité et le fidélité des instruments de mesure;
- la justesse des seuils de réussite rattachés aux instruments de mesure utilisés;
- certaines caractéristiques de l'élève qui auraient pu influencer ses performances: nervosité, maladie, etc.

Placer les résultats de la mesure dans un cadre plus large permet d'obtenir un jugement nuancé. En contrepartie, afin de ne pas être influencé par des préjugés, il faut se donner des règles précises pour formuler un jugement qui s'appuie sur des faits.

Lors de la planification, il faut donc, en plus de choisir une définition de la compétence d'un élève en rapport avec la décision à prendre, établir des règles de jugement où l'on précise les autres facteurs dont il faudra tenir compte pour porter le jugement. Il faut y réfléchir à l'avance et indiquer ce que l'on acceptera de prendre en compte.

B. Comment juger des autres éléments pertinents à l'orientation de l'élève dans sa carrière scolaire?

Dans les décisions qui portent sur le passage à la classe supérieure et le cheminement scolaire de l'élève, on a à porter des jugements qui peuvent être différents de ceux que l'on pose lors de la sanction des études

Exemple de résultats d'élèves

	Objectifs					Somme des objectifs atteints	Moyenne des objectifs atteints
Élèves	**1**	**2**	**3**	**4**	**5**		
Julie	R	R	R	R	E	4/5	0,80
Louis	R	R	E	R	E	3/5	0,60
Martine	E	R	R	E	R	3/5	0,60
Manon	E	R	E	R	R	3/5	0,60
Philippe	E	E	E	E	E	0/5	0,00

E: Échec R: Réussite

ou de la reconnaissance des acquis expérientiels.

On conçoit facilement que le jugement que l'on porte avant de décider qui fera partie d'un programme d'activités pour enfants doués est différent de celui dont on a besoin pour la sanction des études. Dans le cas du placement «enfant doué», on voudra savoir qui est capable de performance supérieure selon des normes établies scientifiquement et dans le cas de la sanction des études, on voudra connaître qui a satisfait aux exigences. S'il s'agit d'accepter un élève dans un cours optionnel au secondaire, on doit prévoir comment on en arrivera à juger de sa préparation à suivre ce cours. Il faut définir ce qu'on appelle les exigences.

En fait, dans tous les cas d'orientation, on a besoin de fixer le poids des prédicteurs par rapport aux critères choisis. Ainsi, pour orienter l'élève dans son cheminement au primaire et au secondaire (passage à la classe suivante ou cheminement particulier), on prévoit comment on jugera de sa compétence par rapport aux préalables pédagogiques. Pour ce faire:

- on identifie les apprentissages vraiment nécessaires à la poursuite des études et les exigences que l'on aura face à ces apprentissages;
- on prévoit comment on jugera du degré de maturité intellectuelle ou psychosociale de l'élève;
- on convient des critères sur lesquels seront basés les jugements.

Ainsi, dans les cas d'orientation, des éléments autres que la compétence en regard des programmes d'études sont objets de jugements. Selon la situation, on devra prévoir comment juger de l'intérêt de l'élève, de sa motivation, de ses attitudes, de sa compétence par rapport à des préalables et même parfois de ses aptitudes par exemple.

La démarche à suivre ressemble à celle qui a été suggérée pour juger de la compétence par rapport à un programme. Il s'agit de:

- définir ce que l'on veut juger (intérêt, attitudes, préalables, etc.);
- décrire les comportements ou les caractéristiques que l'on considérera comme satisfaisants et ceux que l'on considérera insatisfaisants;
- valider cette définition et ces descriptions quant à leur clarté et à leur pertinence auprès de gens compétents en didactique, en psychologie, en pédagogie, etc.); réviser s'il y a lieu la définition et la description;
- choisir les personnes qui porteront les jugements.

C. Comment traiter les cas limites et les cas surprises?

Dans la plupart des cas, les règles générales s'appliquent sans problème. Elles demeurent toutefois plus difficiles à utiliser dans les cas limites et les cas surprises: l'élève qui a un niveau de compétence à peine suffisant, l'élève qui est jugé compétent par son enseignante ou son enseignant mais qui ne réussit pas une épreuve-synthèse uniforme, l'élève qui est dans l'impossibilité de passer l'épreuve de fin d'année, l'élève qui conteste le jugement porté par l'enseignante ou l'enseignant, etc.

Il est conseillé d'essayer de prévoir le plus possible ces cas et d'établir par consensus une façon de faire. Il s'agira souvent de situations où l'on manque d'informations sur lesquelles baser les jugements, où l'on considère que l'information disponible peut être biaisée ou erronée, où les règles de jugement s'avèrent inapplicables. On peut alors penser à prendre en compte d'autres informations, à aller chercher les informations nécessaires selon d'autres moyens (entrevues, travail-

synthèse, etc.), à utiliser des comités où chaque membre explique comment il porte ses jugements ou encore, à accepter de faire une exception à la règle établie pour le jugement. Il s'agit donc d'essayer de prévoir le plus de cas possible pour ne pas être pris au dépourvu.

D. Déterminer les moments où l'on portera les jugements

On conseille toujours de juger de la maîtrise des apprentissages d'un élève le plus près possible de la fin de la période prévue pour cette apprentissage. Ceci pour deux raisons: laisser à l'élève le temps auquel il a droit pour développer les apprentissages au programme et faire en sorte que le jugement reflète le mieux possible la compétence à laquelle l'élève est parvenu. En fait, on fixe les moments de l'évaluation sommative selon la nature des programmes d'études et la répartition du contenu qui est faite.

Par rapport à l'orientation pédagogique, pour plusieurs programmes du secondaire, on pourrait choisir plusieurs moments d'évaluation sommative: un à la fin de chacun des modules qui les composent. Cependant à la fin de l'année, il ne faut pas oublier de se prononcer sur les habiletés transférables d'un module à l'autre. Par ailleurs, en enseignement du français en troisième secondaire, ne serait-il pas logique d'attendre la fin de l'année pour porter un jugement définitif sur le développement des habiletés de compréhension et de production de l'élève? Au primaire, à l'intérieur des cycles eux-mêmes, il y a des objectifs dont la poursuite est terminale à des classes bien précises. Un jugement doit alors être porté à la fin de l'année en question.

L'idée est d'être en mesure de porter un jugement valable sur l'état de l'apprentissage de l'élève en fonction de la décision à prendre. Prenons le cas d'une élève de 4e année primaire. Cette élève a eu des résultats très faibles aux mois d'octobre et de novembre. La pratique de porter un jugement à partir de la moyenne des notes de l'année fera que cette élève échouera quelle que soit sa performance pour le reste de l'année. La question est posée: doit-on la faire doubler l'an prochain parce qu'à l'automne elle n'a pas atteint les objectifs visés?

Ne doit-on pas plutôt porter plusieurs jugements sur son niveau d'apprentissage vers la fin de l'année, sa façon d'apprendre, sa motivation, pour être en mesure de prendre une décision plus juste?

En résumé, on doit planifier l'étape jugement en fonction de la décision à prendre:

— décider sur quoi portera le jugement(exigences établies);
— choisir le type d'informations dont on tiendra compte (liées au programme ou à d'autres facteurs);
— décider des moments où l'on portera les jugements.

Pour assurer la qualité des jugements, il est conseillé de:

— travailler le plus posssible en équipe;
— réviser annuellement les règles de jugement;
— prévoir comment on contrôlera la qualité des informations sur lesquelles on basera ses jugements.

On doit de plus s'efforcer de favoriser la stabilité du jugement que l'on posera. La transparence dans la façon d'établir les règles de jugements facilitera l'acceptation de ces règles.

3.3 Planifier l'étape mesure

Si on veut être en mesure de ne pas se tromper dans nos jugements et nos décisions, il faut se baser sur des informations de qualité. Cette qualité s'obtient à partir des

prises d'informations que l'on fera et de la façon d'analyser et d'interpréter les données.

Il est normal que l'enseignante ou l'enseignant procède à plusieurs prises d'informations pour fonder son jugement. Les méthodes utilisées peuvent être systématiques ou non. Étant donné le type de décision en jeu en évaluation sommative, il est recommandé de procéder de façon systématique quand approche le moment du jugement. Il est même prudent de prendre régulièrement ce qu'on appelle des «portraits» modifiables de l'élève au cas où se produirait une perturbation vers la fin de la période consacrée aux apprentissages visés, pertubation qui empêcherait d'obtenir des informations valables. On tiendra compte de ces différents «portraits» pour tracer le «portrait» final de la compétence de l'élève. Les méthodes systématiques utilisées pour décrire les comportements de l'élève comprennent les observations par l'enseignante ou l'enseignant, les questionnaires, les interviews, les projets de classe, les travaux à long terme et les épreuves écrites.

Il est essentiel que les informations qui seront interprétées soient valables si l'on veut favoriser la stabilité des jugements qui seront portés. Pour cela, les façons de faire (questions d'examens, éléments d'observation...) doivent être adaptées à ce qu'on veut mesurer selon la décision à prendre. Elles doivent également permettre d'identifier la vraie position de l'élève par rapport aux apprentissages visés, à ses aptitudes personnelles, à ses intérêts. On parle alors de **validité.**

Dans le cadre de la sanction des études et de la reconnaissance des acquis expérientiels, on se préoccupe de validité de contenu. On veut que les prises d'informations soient représentatives des domaines qu'elles sont supposées représenter: les apprentissages visés par le programme et donc la compétence recherchée.

Dans le cadre de l'orientation ou du placement, on pourra mettre l'accent sur la validité de construit ou sur la validité prédictive. Ces concepts sont expliqués au tableau 4. Par exemple, à l'aide d'observations ou de tests, on voudra connaître les intérêts des élèves, ou encore on voudra identifier les chances de succès de l'élève qui passe à la classe supérieure.

Comme le temps impose des limites, il est important que l'échantillon d'informations que l'on obtient possède de façon proportionnelle les caractéristiques de l'ensemble visé, le programme de l'année par exemple.

Quel que soit le type de décision en jeu, lors de la planification des prises d'informations, on cherche à décrire d'une façon ordonnée et définie le contenu du domaine en termes de tâches, de comportements.

Voici des questions qui peuvent guider lors de la planification et aider à veiller à la validité des informations recueillies:

— Quelles prises d'informations refléteront les apprentissages visés?

— Comment s'assurer que les items des épreuves auront une bonne sensibilité à l'apprentissage, qualité importante dans l'approche critériée?

— Quelles prises d'informations permettront d'identifier comme très compétents ceux qui sont déjà reconnus comme tels?

— Quelles prises d'informations permettront de prédire le succès dans un domaine relié?

Une caractéristique importante contribue à la validité des prises d'informations, c'est le degré avec lequel elles fournissent des informations constantes(fidélité). En effet, on ne voudrait pas que les résultats à un examen varie d'un jour à l'autre chez le même élève, on ne voudrait pas qu'à deux questions vérifiant le même apprentissage l'élève ait deux

scores différents à moins que l'apprentissage de l'élève ne soit pas stabilisé. Sinon comment pourrait-on se fier à ces informations?

Comment améliorer cette constance, cette fidélité?

— bien clarifier toutes les situations de prise d'informations (préciser les questions, les directives,...);
— améliorer la concordance entre les observatrices et les observateurs ou les différentes occasions d'observation;
— prendre plusieurs informations pour ne pas risquer d'être influencé par des cas exceptionnels (plusieurs questions dans un examen, plusieurs examens, plusieurs éléments et occasions d'observation);
— augmenter la sensibilité des prises d'informations aux différences qu'on veut observer (niveaux d'apprentissage, niveaux de compréhension en lecture).

On pourra ainsi en arriver à une information plus fidèle, sûre et répétable.

A. Choisir la façon d'obtenir les informations nécessaires aux jugements à porter

Lors de la planification du jugement, on a déjà choisi l'ensemble des apprentissages dont on veut tenir compte. Maintenant on décide de la façon dont on ira chercher l'information; on cherche l'instrumentation qui permettra de recueillir l'information sur ces apprentissages. Pour ce faire, on utilise des techniques et des instruments de mesure adaptés à ce qu'on veut mesurer.

Il existe plusieurs types d'épreuves auxquelles on peut avoir recours pour obtenir l'information nécessaire aux jugement à porter, en voici une liste succincte.*

a) épreuve écrite à courtes réponses construites (papier-crayon);
b) épreuve de production écrite(dissertation, rédaction, essai);
c) épreuve écrite à réponses choisies(choix multiples, appariement...);
d) épreuve mixte (réponses construites ou choisies, observation avec liste de pointage, échelle d'appréciation);
e) épreuve orale (langues,...);
f) épreuve pratique (réparation d'un moteur, exécution d'une sculpture, exécution d'une improvisation en danse...);
g) épreuve à longue échéance (travail de recherche, expérience...);
h) examen qui combine plusieurs épreuves.

Le choix du type d'épreuve doit reposer sur le type d'apprentissage à observer. Par exemple, il est évident qu'on ne mesurera pas le développement d'habiletés psychomotrices par une épreuve papier-crayon; sauf dans certains cas pour la motricité fine.

Il s'agit donc à cette étape-ci d'analyser les objectifs d'apprentissage retenus pour l'évaluation afin d'en identifier les caractéristiques et de spécifier les tâches évaluatives ou les épreuves nécessaires. Plusieurs méthodes sont disponibles. Pour l'analyse des objectifs d'apprentissage, nous présentons l'analyse taxonomique et l'analyse par facettes. Nous présentons également deux méthodes pour spécifier les tâches évaluatives: l'analyse selon le centre IOX (Instructional Objectives Exchange) et la forme d'item. Dans les pages suivantes un exemple de chacune de ces méthodes est présenté.

- **L'analyse taxonomique**: on analyse les apprentissages visés selon le type de comportements intellectuels, affectifs, psychomoteur ou le type de produits d'apprentissage. Les catégories d'une taxonomie peuvent servir à construire le tableau de spécifications de l'épreuve. Il

* Pour plus de renseignements voir le fascicule 5 des *Éléments de docimologie* (à paraître).

existe plusieurs taxonomies comme celles de Bloom, de Gagné, de Harrow.

Par exemple, dans la taxonomie de Harrow, les niveaux de classification sont situés dans une hiérarchie. On y décrit un continuum qui va du plus bas niveau de comportement moteur observable au plus haut niveau. L'utilité d'une taxonomie réside dans le fait que l'analyse des apprentissages visés en fait ressortir les caractéristiques et que la connaissance de ces caractéristiques favorise l'élaboration ou le choix d'items adéquats.

La taxonomie de Harrow

1.00 Les mouvements réflexes
- 1.10 Les réflexes segmentaires
- 1.20 Les réflexes plurisegmentaires
- 1.30 Les réflexes suprasegmentaires

2.00 Les mouvements fondamentaux
- 2.10 Les mouvements locomoteurs
- 2.20 Les mouvements non-locomoteurs
- 2.30 Les mouvements de manipulation

3.00 Les capacités perceptives
- 3.10 La discrimination kinesthésique
- 3.20 La discrimination visuelle
- 3.30 La discrimination auditive
- 3.40 La discrimination tactile
- 3.50 La coordination

4.00 Les capacités physiques
- 4.10 L'endurance
- 4.20 La force
- 4.30 La souplesse
- 4.40 L'agilité

5.00 Les habiletés motrices
- 5.10 d'adaptation
- 5.20 d'adaptation composite
- 5.30 d'adaptation complexe

6.00 La communication gestuelle
- 6.10 Les mouvements d'expression
- 6.20 Les mouvements d'interprétation.

- **L'analyse par facettes:** on analyse les apprentissages pour en identifier les composantes ou les aspects importants qu'on appelle facettes. Une combinaison de facettes représente une dimension de l'apprentissage visé. Celui-ci comprend donc autant de dimensions qu'il y a de combinaisons possibles. Ainsi, dans l'exemple présenté on a identifié deux facettes: l'intention de lecture et l'opération mentale faite par la lectrice ou le lecteur. Une dimension serait la combi-

Exemple d'analyse par facettes – lecture, troisième année du primaire

Intentions de lecture

1. Prendre une décision pratique.
2. Se renseigner et éveiller ou satisfaire certaines curiosités.
3. Dégager une conduite à adopter ou une règle pour agir.
4. Résoudre un problème.
5. Se divertir.

Opérations mentales

1. Le **repérage**: l'information dont l'élève a besoin pour exécuter sa tâche est donnée intégralement dans le texte. L'élève n'a qu'à repérer les mots, les phrases qui se rattachent explicitement à l'information cherchée.

2. La **sélection**: l'information dont l'élève a besoin pour exécuter sa tâche est donnée dans le texte parmi plusieurs informations sur le même sujet qui n'ont pas trait à ce qu'il cherche. L'élève doit alors sélectionner parmi les informations données celles dont il a besoin de celles qui sont non pertinentes pour lui selon son intention.

3. Le **regroupement**: les informations dont la lectrice ou le lecteur a besoin pour exécuter sa tâche sont dispersées à plus d'un endroit dans le texte. L'élève doit alors repérer ces informations, souvent les sélectionner pour établir des liens entre elles, les regrouper, soit pour les résumer tout simplement, les comparer, les confronter ou les juger.

4. L'**inférence**: certaines informations dont la lectrice ou le lecteur a besoin pour exécuter sa tâche ne sont pas données explicitement dans le texte. Pour les reconnaître, cette personne doit faire appel à des connaissances qu'elle a déjà ou à des informatons déjà données explicitement dans le texte: elle doit alors inférer.

5. Le **jugement**: l'intention de lecture peut inciter la lectrice ou le lecteur à juger des informations qu'elle ou qu'il lit soit pour distinguer un fait d'une opinion ou d'un sentiment, soit pour distinguer le réel de l'imaginaire ou du fictif.

Note: les opérations telles l'inférence et le jugement d'informations ne devraient pas faire partie des épreuves-synthèse en troisième année.

Inspiré de Français langue maternelle au primaire. *Définition de domaine* – 3e année. Monique Careau, C.S. Montcalm, et Nicole Leblanc, C.S. des Îlets. Version provisoire, 1983.

naison d'une intention de lecture (se renseigner) et d'une opération mentale (la sélection). À cette dimension on pourrait faire correspondre la tâche évaluative suivante: lire avec l'intention de se renseigner un dépliant donnant de l'information sur la bicyclette.

- **L'analyse selon le centre IOX** (Instructional Objectives Exchange): selon ce modèle, la tâche évaluative est analysée par rapport à cinq éléments: une description générale de la tâche, un exemple d'item, les caractéristiques du stimulus (la question), les caractéristiques de la réponse et des instructions supplémentaires s'il y a lieu. L'item donné en exemple ci-dessous porte sur le choix d'un titre pour un texte en français.
- **La forme d'item:** c'est une autre façon d'analyser la tâche évaluative à partir d'une liste de règles permettant de produire une ensemble d'items à structure syntaxique fixe. On énonce d'abord l'objectif qui présente le comportement et le contenu. On décrit de façon générale les caractéristiques de ce contenu et de ce comportement. On décrit aussi le stimulus. On prépare une matrice cellulaire pour présenter l'ensemble des stimuli possibles, c'est l'étendue du domaine. À

Exemple d'analyse selon le centre IOX

Description générale. À partir d'une histoire, l'élève devra identifier parmi 4 titres possibles, celui qui lui convient le mieux.

Exemple d'item. Consigne: Lis le texte suivant et réponds aux questions en encerclant le numéro placé devant ta réponse.

Il était 5 heures de l'après-midi quand notre famille revient de la plage. La journée avait été resplendissante; nous étions partis tôt dans l'avant-midi; le trajet de la maison à la plage était long et nous voulions arriver pour le dîner. De plus, il fallait prendre chez lui un de mes copains qui allait nous accompagner.

Nous étions à peine sortis de l'auto que notre voisin de droite arriva en courant et en gesticulant. Il s'approcha de mes parents, les amena quelque peu à l'écart. Visiblement, il ne voulait pas que j'entende. La conversation s'anima. Cela devenait intrigant. Lentement, je m'approchai. Ils ne me remarquèrent pas, trop préoccupés qu'ils étaient à discuter. Ce que j'entendis alors ne m'a guère réjoui. Il venait d'être cambriolé et croyait que c'était un de mes copains. – Je m'avançai un peu plus car j'entendais mal. – Il l'aurait vu flâner dans le voisinage tout l'avant-midi. Vers l'heure du midi, alors qu'il était allé se promener, le copain en aurait profité pour pénétrer chez lui. Un autre voisin, le voyant faire, aurait appelé la police qui serait arrivée trop tard, le malfaisant ayant déjà pris la fuite.

Tous mes copains ne sont pas des anges, pensai-je, mais aucun n'est capable de poser un tel geste. Je lui demandai le nom de cette personne. Il l'ignorait. Levant alors les yeux, il se mit à crier: «Le voilà! C'est lui, je le reconnais ! Au voleur!»

Nous étions tous un peu gênés de le voir ainsi s'agiter en parlant de mon copain qui venait remercier à nouveau mes parents pour la magnifique journée qu'il avait passée avec nous à la plage.

Questions:

1. Quel titre convient le mieux à cette histoire?
   A) Une ténébreuse affaire.
   B) Une journée à la plage.
   C) Erreur sur la personne.
   D) Mon meilleur copain.

Caractéristiques du stimulus. Il s'agit d'un texte fictif, complet en lui-même, ne contenant pas plus de trois cents mots et pas moins de cent. Les mots doivent être simples, tirés du vocabulaire de l'élève. Le texte ne devra pas comporter de titre. Etc., etc.

Caractéristiques de la réponse. Les choix sont au nombre de 4. Ils ne doivent pas comporter de verbe: il s'agit de syntagmes nominaux. Pour la bonne réponse, il n'est pas nécessaire que les noms qui forment ces syntagmes soient tirés du texte. Il n'y a qu'une seule bonne réponse par question. La forme des leurres doit ressembler à celle de la bonne réponse (ex. être sensiblement de la même longueur). Deux de ces leurres doivent contenir au moins un nom tiré du texte. L'élève devra encercler le chiffre (ou la lettre selon le cas) placé devant son choix. Si une feuille-réponse est fournie, il devra alors inscrire ce chiffre devant le numéro de sa réponse.

chaque cellule, correspond un stimulus possible que l'on décrit dans la section «Composition des cellules». Ainsi dans l'exemple suivant, la cellule «A» indique qu'on présentera à l'élève deux ensembles égaux de quatre éléments chacun; les éléments d'un ensemble seront bleus et ceux de l'autre seront jaunes. On fournit aussi les directives que l'enseignante ou l'enseignant devra suivre au moment de la passation de l'épreuve et au moment de la correction des réponses des élèves. On peut aussi inclure une section consacrée à la description du matériel utilisé. Cette description doit être la plus précise possible pour ne pas laisser place à l'interprétation.

B. Choisir les façons d'interpréter les informations et d'établir le seuil de réussite s'il a lieu

Planifier l'élaboration des épreuves n'est pas suffisant. En plus, il faut en arriver à être en mesure d'interpréter le résultat de l'individu. Savoir qu'un élève a 15 bonnes réponses sur 20 ne nous avancera pas beaucoup. On a besoin de savoir si ce résultat correspond aux exigences préalablement définies. Il nous faut alors un point de référence,

Exemple d'une forme d'item

L'élève sera en mesure de comparer le nombre d'éléments que contiennent des ensembles en faisant correspondre un à un les éléments de chacun des ensembles.

a) Description générale: On présente à l'enfant deux ou trois ensembles contenant moins de 17 jetons. Les ensembles ne doivent pas différer entre eux de plus d'un jeton. On demande à l'enfant de dire, sans compter, si les ensembles qu'on lui présente comportent le même nombre d'éléments.

b) Stimulus: On utilise des jetons de taille approximativement égale à celle d'une pièce de 5 cents. Les ensembles se distinguent par la couleur des jetons qu'ils contiennent.
On peut demander à l'élève de comparer soit 2 ou 3 ensembles de jetons. Le nombre de jetons peut être différent d'un ensemble à l'autre. Les ensembles doivent contenir approximativement 4, 9 ou 16 jetons.

c) Matrices cellulaires

Nombre d'éléments	Nombre d'ensembles à comparer				
	2		3		
	Type de relation d'égalité		Type de relation approximatif d'égalité		
	Na≠Nb	Na≈Nb	Na=Nb=Nc	Na=Nb≠Nc	Na≈Nb≈Nc
4	A	D	G	J	M
9	B	E	H	K	N
15	C	F	I	L	O

(Na ≠ Nb: Les ensembles de jetons bleus et jaunes n'ont pas le même nombre d'éléments.)

d) Composition des cellules

• Couleur des jetons:
bleu=a jaune=b vert=c

Cellule A: 4a-4b
Celule B: 9a-9b
Cellule C: 15a-15b
Cellule D: 4a-5b
Cellule E: 9a-10b
Cellule F: 15a-16b
Cellule G: 4a-4b-4c
Cellule H: 9a-9b-9c
Cellule I: 16a-16b-16c
Cellule J: 4a-4b-5c
Cellule K: 9a-9b-10c
Cellule L: 15a-15b-16c
Cellule M: 4a-5b-6c
Cellule N: 9a-10b-11c
Cellule O: 14a-15b-16c

e) Matériel: Jetons bleus, jaunes et verts pour former des ensembles.

f) Directives: 1. Placer les ensembles côte à côte, ensuite dire à l'élève:
«Dis-moi si ces ensembles (2 ou 3) ont le même nombre d'éléments.»

• *Si* l'élève se met à compter ou dit ne pas savoir, alors lui dire:
«En classe, on a déjà pairé des objets pour savoir si deux ensembles possédaient le même nombre d'éléments. Maintenant, dis-moi si ces ensembles (2 ou 3) ont le même nombre d'éléments.»

• Lorsque l'élève a fini de pairer les jetons, lui dire:
«Ont-ils le même nombre d'éléments?»

2. Noter les commentaires et actions de l'élève.

g) Règles de corrections: L'élève doit dire clairement si oui ou non les ensembles contiennent le même nombre d'éléments et cela en alignant les jetons.

Inspiré de *Six modèles de spécifications, Complément au Guide pour la construction d'instruments de mesure dans le cadre de la Banque d'instruments de mesure, 16-7300-07.*

un point de comparaison, pour être capable d'interpréter le score à une épreuve.

Il y a deux types de points de référence. De façon stricte, une épreuve à référence normative a déjà été expérimentée et on possède de l'information sur le groupe spécifique d'individus qui l'ont passée (scores standardisés, stanines, percentiles). Une épreuve à référence critériée est directement reliée à des objectifs des tâches, des éléments de compétence. Elle peut être interprétée en termes d'acquisition des habiletés et des connaissances visées par le programme, donc par rapport aux comportements attendus. Les résultats de l'élève reflètent son répertoire de comportements, c'est-à-dire ce qu'il est capable de faire et ne dépendent pas de la performance des autres. Pour établir le seuil de réussite, on se demande quel est le score qui discriminerait ceux qui sont compétents de ceux qui ne le sont pas par rapport aux habiletés visées.

Mais qu'est-ce qu'un **seuil de réussite**? C'est **le niveau de qualité à partir duquel on considère une performance comme réussie**, c'est-à-dire comme reflétant la compétence. On peut définir la compétence comme étant la capacité à utiliser un comportement ou un ensemble de comportements pour satisfaire les besoins d'une situation. C'est donc la mesure de ce qui est adéquat pour un certain but. Dans le contexte où l'on retrouve un seuil de réussite relié à un instrument de mesure, ce dernier étant lui-même relié à une définition du domaine, le seuil de réussite peut se définir comme un score qui divise une distribution de scores en deux parties distinctes: «la première contenant les scores à partir desquels on infère la compétence des individus et la seconde, les scores à partir desquels on infère l'atteinte d'un degré insuffisant de compétence».*

* *Guide pour l'établissement du seuil de réussite*, Réjean Auger, mars 1986, 16-7514, p. 2.

Toute méthode utilisée pour établir ce seuil comprend un jugement et implique un certain degré de variation. On doit éliminer les facteurs non pertinents de telle sorte que les jugements soient aussi objectifs que possible donc, reproductibles. De toute façon, il ne faut jamais considérer un seuil de réussite comme étant un nombre magique. En fait, il y a toujours une marge d'incertitude autour du score qui est choisi comme seuil de réussite.

Le contexte dans lequel se situera l'utilisation du seuil de réussite doit être pris en compte. Lorsque l'on établit un seuil de réussite, il faut considérer l'importance de la décision à prendre. Le risque d'erreur est ici encore présent. Quoiqu'il soit très difficile de répondre à la question, on peut se demander s'il est plus grave de faire passer un élève qui ne maîtrise pas l'apprentissage que de faire échouer un élève qui le maîtrise. D'autres questions peuvent être soulevées, à savoir par exemple si les élèves ont été suffisamment en contact avec les objets d'apprentissage pour développer l'habileté visée, s'ils auront l'occasion de reprendre l'épreuve. On rencontre à nouveau, ici, le risque de commettre des erreurs. Il ne s'agit donc pas seulement d'un problème statistique mais aussi d'un problème à portée pédagogique, psychologique et même sociale. On peut avoir en tête une certaine conception des exigences en rapport à un apprentissage. Mais ce n'est généralement pas assez précis pour servir de base à une règle objective de décision.

C'est pourquoi il faut arriver par consensus à exprimer ces exigences sous forme d'une information observable.

Il existe plusieurs méthodes pour établir les seuils de réussite (Angoff, Nedelsky, Ebel, Block...). Dans chacune on cherche à dichotomiser, à séparer en deux des habiletés qui sont continues. On n'est pas habile ou inhabile mais plus ou moins habile. Ce point

de coupure qu'on veut établir est artificiel et aucune méthode n'est plus valide ou logique qu'une autre pour l'établir. Malgré tout, ces méthodes demeurent des stratégies valables et pratiques pour réfléchir à ce que devrait être le seuil de réussite et pour en fixer un.

Quelle que soit la méthode utilisée, il semble que les considérations suivantes s'appliquent:

a) La notion de seuil se réfère à une décision dichotomique (oui/non, maîtrise/non-maîtrise) tandis que les habiletés humaines ne sont pas distribuées de cette façon.

b) L'établissement du seuil de réussite ne doit pas être seulement une démarche scientifique. On doit tenir compte des systèmes de valeurs. Tout seuil est essentiellement le résultat d'un choix et des données empiriques permettent d'avoir des discussions mieux informées et plus réalistes. Le jugement est à la base de l'établissement du seuil. Le choix des juges, leur expérience et leur sagesse est d'une importance cruciale. Mais tout groupe de juges peut se tromper.

c) Aucune technique d'établissement de seuil n'est suffisante en elle-même, mais ces techniques nous aident. Toutefois, il n'existe pas en soi de méthode meilleure qu'une autre.

d) L'établissement du seuil de réussite est une responsabilité qui implique des conséquences.

e) L'importance accordée à chaque type d'erreur varie selon le contexte et le type de décision à prendre. Aucune mesure unique (une seule épreuve) ne peut refléter l'étendue des compétences et des caractéristiques qui sont minimales. L'utilisation exclusive d'épreuves pour porter le jugement de compétence n'est pas suffisante, on doit utiliser d'autres données telles que la valeur des épreuves, les circonstances de passation de l'épreuve, le vécu de l'élève, etc.

Quelles sont alors les caractéristiques de la méthode la mieux appropriée pour établir les seuils de réussite?

a) La méthode doit tenir compte du but visé par l'utilisation de l'instrument de mesure. Sa valeur technique doit la rendre adéquate au but visé. Par exemple, elle doit conduire à une classification en termes de maîtrise ou de non-maîtrise, elle doit permettre de fixer un seuil de réussite qui reflète la performance actuelle des élèves, elle doit utiliser un instrument congruent avec l'apprentissage vécu par les élèves, elle doit utiliser les bonnes statistiques pour résumer ou interpréter les jugements des experts, elle doit permettre d'estimer le seuil de réussite en tenant compte de l'erreur de mesure (zone d'incertitude).

b) La méthode doit être facile d'application, de compilation et d'interprétation.

c) La méthode doit être adaptée aux ressources disponibles: temps, argent, nombre de personnes et capacité de ces personnes. Ces dernières doivent être qualifiées et bien comprendre la méthode.

d) La méthode doit être adaptée au moment choisi pour fixer le seuil de réussite, c'est-à-dire avant ou après une première administration de l'instrument de mesure.

e) La méthode doit tenir compte des conséquences des deux types d'erreurs: identifier comme incompétent quelqu'un de compétent et identifier comme compétent quelqu'un d'incompétent.

De plus, on doit prévoir réviser régulièrement les seuils de réussite car le contexte change continuellement et notre connaissance de ces seuils s'approfondit à mesure qu'on les emploie.

La tâche sera plus difficile si l'apprentissage visé et l'épreuve ne sont pas spécifiés précisément. De plus, on peut s'attendre à ce que le seuil de réussite soit différent d'une épreuve à l'autre et selon le contexte d'apprentissage.

Il est recommandé de tenir compte de jugements absolus et relatifs. Jugements absolus, c'est-à-dire considérant seulement l'exigence par rapport aux apprentissages visés, ceci pour éviter de déclarer compétents des élèves franchement incompétents. Jugements relatifs, c'est-à-dire condidérant seulement la proportion acceptable d'échecs, pour éviter des exigences irréalistes.

Lors de la planification, on prévoit une période d'une durée suffisante pour établir le seuil de réussite. Pour ce faire, on recule dans le temps à partir de la date où la démarche doit être complétée en considérant le temps requis pour chacune des activités suivantes:

— sélectionner et entraîner les juges;

— si on utilise une nouvelle méthode, l'essayer à une petite échelle avant;

— si on fait passer l'épreuve, l'imprimer, l'administrer, la corriger.

Il faut toujours allouer plus de temps que ça ne paraît nécessaire parce que ça peut être difficile de trouver des moments où tous les juges sont libres. Il se peut aussi qu'après avoir recueilli les jugments, on s'apperçoive que certains juges n'ont pas bien compris la tâche et qu'on doive répéter l'opération. Il faut donc se garder une marge de manoeuvre.

La façon de faire aura un grand impact sur l'acceptation des seuils fixés. Plus la méthode sera connue et ouverte, mieux les seuils seront acceptés.

Pour contrôler la variation entre les opinions des juges, il est conseillé:

— de choisir des personnes qui connaissent bien l'habileté sur laquelle porte l'épreuve;

— d'avoir un nombre suffisant de juges (5 à 8 juges compétents);

— de bien animer et diriger les divers groupes de travail.

Méthodes

On peut regrouper en deux grandes catégories les méthodes pour établir le seuil de réussite selon qu'elles recourent à des jugements sur le contenu de l'épreuve ou sur les élèves qui ont passé l'épreuve.

Seuils basés sur les jugements du contenu de l'épreuve: dans ces méthodes, on établit un seuil absolu en analysant le contenu de l'épreuve. En se basant sur chaque question, on juge de la probabilité qu'à un élève de compétence minimale de réussir. On se représente l'image de ce qu'est la compétence minimale à partir de l'instrument de mesure. On retrouve parmi ces méthodes celles de Angoff, Nedelsky, Ebel... À titre indicatif, on présente succinctement celle de Angoff. Pour un exposé détaillé de ces méthodes, on peut consulter le fascicule 5 des *Éléments de docimologie.*

La méthode de Angoff. Les juges se représentent l'élève qui a une compétence minimale en regard de l'épreuve. Il est nécessaire d'avoir une image mentale de ce type d'élève. Par la suite, les juges attribuent une probabilité à chaque item. C'est la probabilité qu'un élève de compétence minimale réponde correctement à l'item. On calcule pour chaque juge la probabilité de réussite à l'épreuve. Par la suite, on établit le seuil de réussite en calculant cette probabilité pour l'ensemble des juges. Cette méthode est simple et permet aux juges d'être conscients de la subjectivité de leurs déclarations. La validité de ce type de méthode dépend de la définition de l'élève qui a une compétence minimale. Les juges

doivent être entraînés et avoir l'occasion de discuter ensemble. En ne tenant pas compte des résultats de l'ensemble des élèves, les juges peuvent identifier un seuil de réussite absolu.

Toutes les méthodes basées sur des jugements à propos des questions d'une épreuve comportent six étapes:

1. Sélectionner les juges.
2. Se représenter la performance minimale acceptable.
3. Entraîner les juges à utiliser la méthode choisie.
4. Permettre la discussion entre les juges.
5. Recueillir les jugements.
6. Combiner les jugements pour établir le seuil.

Ces méthodes peuvent être appliquées avant ou après l'administration de l'épreuve. Se situant dans une démarche de prise de décision, elles demandent à être confrontées à la réalité. On vérifie alors, après que le seuil ait été fixé sans connaître les résultats des élèves, s'il correspond à la performance de ceux d'entre eux qui ont une compétence minimale. En effet, le jugement à propos de ce que les élèves devraient être capables de faire dépend d'une certaine façon de ce qu'ils peuvent faire. La limite de ces méthodes est que les jugements peuvent être inconsistants et même tout simplement irréalistes.

Seuils basés sur les jugements à propos des élèves qui ont passé l'épreuve: on a ici besoin de deux types d'informations à propos de chaque élève: son résultat à l'épreuve et un jugement de sa compétence. Ces méthodes comprennent la méthode du «groupe frontalier», celle des «groupes opposés» et une variation de cette dernière, celle du «tâtonnement». L'avantage de ces méthodes est qu'elles ont recours à un procédé familier, c'est-à-dire celui de juger si l'habileté d'une personne est suffisante dans un contexte particulier.

Les jugements doivent remplir les exigences suivantes:

1. Ils doivent être faits par des personnes qualifiées pour les faire (bonne connaissance de l'épreuve et de la compétence de chaque élève).
2. Ils doivent porter sur les connaissances et sur les habiletés mesurées par l'épreuve.
3. Ils doivent refléter les habiletés des élèves au moment de la passation de l'épreuve.
4. Ils doivent refléter l'opinion véritable des juges.

Dans tous les cas, avant de porter leur jugement, les juges ne doivent pas connaître les scores obtenus par les élèves.

Voici la description succinte de la méthode du «**groupe frontalier**»*. Cette méthode est basée sur l'idée que le seuil de réussite est le score que l'élève de compétence minimale devrait obtenir. Pour appliquer cette méthode, les juges identifient les élèves qu'on ne peut qualifier, avec certitude, de compétents ou d'incompétents. Le seuil de réussite est le score milieu (la médiane) des scores de ce groupe d'élèves. L'avantage de cette méthode est qu'elle est facile à comprendre et à utiliser. Son grand désavantage est que le nombre d'élèves à compétence minimale ne représente qu'un petit pourcentage des élèves et qu'il est difficile pour les juges de cerner le groupe frontalier.

Dans tous les cas, on choisit la méthode d'établissement du seuil de réusssite en fonction du type de déclaration qui peut le mieux être porté par les experts que l'on prévoit utiliser. Il est recommandé, si c'est possible, d'employer une méthode qui utilise les notes

* Pour un exposé détaillé des autres méthodes on peut consulter le fascicule 5 des *Éléments de docimologie*.

réelles des élèves (méthode a posteriori). Il semble donc que, lors de la planification, on ait à choisir la méthode à utiliser et, s'il y a lieu, qu'on doive décider de fixer un seuil de réussite a priori.

Il est entendu qu'une enseignante ou qu'un enseignant n'a pas beaucoup le temps ni les ressources pour suivre des techniques élaborées. Elle ou il peut alors avoir en tête deux questions qui peuvent le guider: À quelle production minimale acceptable doit-on s'attendre sur tel sujet? Quelle est la performance habituelle des élèves qui réussissent telle classe? Si on a l'impression que trop d'élèves compétents n'atteignent pas le seuil de réussite ou que trop d'élèves incompétents réussissent, on doit en chercher la cause. Ce peut être parce que le seuil de réussite n'est pas adéquat, la méthode n'était pas adaptée, les juges étaient d'opinions très divergentes ou n'étaient pas compétents pour appliquer cette méthode, les élèves ne sont pas aussi compétents qu'on le pensait, l'instrument de mesure est à remettre en question.

Comment interpréter? L'utilisation d'un seuil de réussite fournit un point de repère fort utile. Ce seuil de réussite peut être fixé pour chaque objectif terminal retenu pour fins de décision. Pour chacun des objectifs, on prévoit comment organiser les données pour en faciliter l'interprétation. Ainsi on vise à avoir une information précise sur chacun des critères. Ces renseignements peuvent être inscrits sur la fiche de correction. On compare alors la performance de l'élève aux différents seuils fixés. Il est possible alors au moment du jugement d'établir un profil des apprentissages de l'élève.

Il s'agit donc de traduire en termes d'apprentissages les différentes informations recueillies pour en arriver à dire quel degré de compétence reflètent les performances de l'élève.

N'oublions pas toutefois que le résultat obtenu par l'élève ne représente pas parfaitement son habileté. Trop de sources d'erreurs influencent ce résultat: la formulation de la question, le degré de fatigue de l'élève, la façon de corriger, etc. Il y a donc toujours une marge d'erreur, un intervalle autour du score obtenu où l'on est incertain du score précis qui représenterait exactement l'habileté de l'élève. C'est pourquoi il faut prévoir plusieurs prises d'informations. L'idée est toujours d'essayer d'éviter les deux types d'erreurs: faire échouer l'élève compétent et faire passer l'élève incompétent.

C. Déterminer les moments où l'on procédera à la mesure

La planification de la mesure en évaluation sommative comporte une série de choix qui doivent être faits dans une séquence déterminée. Cela exige la participation de plusieurs personnes. Par conséquent, la planification doit se faire longtemps à l'avance. Voici une liste des choix à faire:

— choisir le type d'examens en fonction de chaque apprentissage retenu:
  - selon le type décision à prendre,
  - selon l'objectif d'apprentissage que l'on a analysé;

— prévoir les façons d'interpréter les données:
  - choisir le point de référence (absolu ou relatif),
  - choisir la méthode d'établissement du seuil de réussite et en planifier l'application,
  - prévoir l'organisation des données qui facilitera l'interprétation.

L'établissement de l'échéancier doit tenir compte du calendrier déjà fixé lors de la planification des décisions et des jugements. Il est prudent de prévoir plusieurs occasions

(questions, épreuves, moments) afin d'avoir pour chaque élève le portrait le plus précis possible des aspects à évaluer. L'enseignante ou l'enseignant doit tenir compte de cette planification dans celle de son enseignement.

3.4 Choisir une façon de conserver les données

En évaluation sommative, la planification s'étend jusqu'à la prévision de la façon de conserver les données. On peut procéder de différentes façons:

— conserver les épreuves, les travaux des élèves;

— conserver les interprétations portées à partir de chaque épreuve d'un ensemble en vue de faire le bulletin (voir l'exemple de la page 56);

— choisir le bulletin que l'on utilisera (traditionnel, descriptif).

À proprement parler, le bulletin n'est pas un élément essentiel de la démarche d'évaluation puisqu'on pourrait très bien prendre des informations sur les apprentissages de l'élève, juger de sa compétence et prendre des décisions à son sujet sans utiliser de bulletin. Toutefois, le bulletin s'inscrit très bien dans le contexte de l'évaluation des apprentissages. Le rôle qu'il y remplit est de renseigner les parents, le personnel enseignant et les institutions qui reçoivent l'élève en début ou en cours d'année. C'est donc dire qu'il s'adresse à une variété de personnes susceptibles d'avoir des besoins d'information différents. On doit tenir compte de cette réalité et choisir un bulletin dont les caractéristiques permettent de fournir des renseignements en nombre suffisant et sous une forme compréhensible pour les personnes qui le reçoivent.

Lors du choix du bulletin, on précise le type d'information que l'on veut y inscrire. Comme on se situe en évaluation sommative, ce qui sera consigné n'est pas le cumul des résultats de l'évaluation formative (devoirs, exercices, etc.). À la fin du cours, on indique le jugement final qui aura été porté le plus près possible de la fin du cours. On peut aussi inscrire la décision prise. Nous avons vu qu'il est prudent de procéder à des évaluations sommatives transitoires ou temporaires pendant le cours. Les jugements temporaires que l'on consigne alors dans les bulletins émis en cours d'année peuvent porter sur des parties de programme. Selon le programme d'études concerné et selon les règles de jugement et de décision que l'on s'est données, la poursuite de certains apprentissages peut être restreinte à une période fixe en cours d'année après laquelle on ne compte plus accorder de temps à ces apprentissages. Dans ce cas, ces jugements auraient un caractère définitif; il serait équitable par ailleurs de se réserver la possibilité de réviser de tels jugements lorsqu'on se rend compte qu'ils ne sont plus exacts. Il faut être bien conscient que ce ne sont pas les échéances du bulletin qui doivent conditionner a priori le déroulement de l'évaluation sommative ni, par la suite, le jugement final et la décision définitive que l'on prend. Ce sont plutôt les règles de jugement et de décision que l'on a fixées précédemment en tenant compte des décisions à prendre.

Il existe plusieurs genres de bulletins. Certains servent à donner une information assez détaillée, d'autres une information plus condensée. On constate que l'utilisation du bulletin descriptif se généralise de plus en plus. Pour qui veut faire un tel choix, il est avantageux de consulter les prototypes de bulletin du ministère de l'Éducation et ceux des commissions scolaires ou des institutions qui en ont élaborés. Le genre de bulletin que l'on choisit reflète les choix pédagogiques et administratifs que l'on fait.

Tout bulletin comporte trois parties. Une section rassemble les renseignements sur l'élève (nom, adresse, etc.). Une autre partie concerne les apprentissages au programme de différentes matières. Une dernière section se rapporte aux comportements et aux attitudes de l'élève. En effet, les données relatives au rendement et au comportement doivent être consignées séparément au bulletin scolaire. Pour chaque matière, on choisit des indicateurs ou descripteurs représentant les parties importantes du programme. On détermine aussi une légende à l'aide de laquelle on communiquera les jugements temporaires ou définitifs sur l'ensemble du programme ou ses parties importantes. Cette légende doit permettre d'exprimer clairement le jugement porté sur la compétence de l'élève.

Lors de la planification, on établit les échéances du bulletin en rapport avec les moments où l'on a décidé que l'on portera les jugements temporaires et le jugement final. Ces échéances sont nécessairement reliées à la structure des programmes d'études et à la nécessité de fournir régulièrement des informations aux parents.

Nous avons vu que le jugement définitif sur la compétence ou l'incompétence d'un élève doit se baser sur plusieurs prises d'informations. Ceci dans le but d'éviter le plus possible les erreurs dans le jugement et la décision qui le suivra. De plus, il est prudent que l'on décrive à intervalle régulier l'état de l'apprentissage de l'élève au cas où il déménagerait ou serait dans l'impossibilité de passer les derniers examens à cause de maladie, de mortalité dans sa famille, d'accident, etc., ou que l'enseignante ou l'enseignant lui-même ne puisse porter ce jugement à cause de maladie, d'accident... Il faut donc prévoir conserver les informations que l'on prend de même que les jugements que l'on porte. On peut consigner les résultats de plusieurs prises d'informations selon les catégories choisies lors de l'établissement des règles de jugement, cette façon de faire doit aider à porter un jugement éclairé.

Les tableaux que l'on utilise alors peuvent être plus ou moins détaillés et ont pour but de faciliter le jugement. L'exemple suivant montre une compilation de résultats de plusieurs épreuves en mathématique 6^{e} année du primaire.

Il est à remarquer que lors de telles compilations il ne faut pas perdre les renseignements fournis aux sections REMARQUES. Ces informations peuvent être d'une importance capitale lors du jugement et de la décision finale.

On pourrait aussi avoir pour chaque élève la liste des objectifs atteints comme dans l'exemple suivant.

Exemple d'un tableau pour enregistrer les interprétations pour *plusieurs épreuves* en mathématique, 6e année du primaire.

NOM: Carole

Habiletés	**A – Nombres**	**B – Géométrie**	**C – Mesure**
1. Discriminer	Ex: réussite de 3 épreuves ①, ②, 3	Ex: réussite de 2 épreuves X ①, ②, 3	Ex: réussite de 2 épreuves X 1, ②, 3, ④
2. Classifier	Ex: réussite de 3 épreuves X ①, 2, ④, ⑤	Ex: réussite de 3 épreuves ①, ②, 3	Ex: réussite d'une épreuve X 1, ②
3. Construire	Ex: réussite de 3 épreuves X ①, ②, ⑤	Ex: réussite de 3 épreuves ①, ②, 5	
4. Estimer	Ex: réussite de 3 épreuves X ①, ②, 3, ④		Ex: réussite de 2 épreuves X ①, ④
5. Opérer	Ex: réussite de 3 épreuves 1, 3, ④	Ex: réussite de 2 épreuves ②, 3	Ex: réussite de 2 épreuves ②, 3

○: Épreuve réussie
Ex: Nombre exigé d'épreuves réussies
X: Dimension réussie

Remarque: *Carole revenait d'une absence de maladie le jour de l'épreuve 3. Je décide de ne pas tenir compte de cette épreuve. Il faut aller vérifier les dimensions 1-A, 2-B, 5-B et 5-C avant de porter le jugement.*

3.5 Résumé

La planification de l'évaluation sommative comporte de nombreuses étapes qui peuvent être assez longues à remplir, compte tenu de l'ampleur des mesures à effectuer, des jugements à porter et des décisions à prendre. Meilleure sera la planification, meilleures seront les actions posées par la suite.

Quand on prépare ce type d'évaluation, on se pose plusieurs questions:

— Quelles décisions puis-je prendre?
— Comment vais-je prendre chaque décision (établir une règle)?
— Comment vais-je porter chaque jugement (établir une règle)?
— De quelles informations ai-je besoin?
— Comment vais-je recueillir cette information (type d'épreuves, de questions...)?
— Quelles tâches vais-je donner à l'élève (travail, questions, problèmes...)?
— Comment vais-je corriger?
— Comment vais-je interpréter les résultats (seuils de réussite, pistes d'interprétation)?

— Quand vais-je poser chaque action?
— Comment vais-je conserver ces données?
— Quel bulletin vais-je choisir?
— Quelles seront les échéances des bulletins?

Lorsque l'on prévoit l'échéancier, il est toujours préférable de se garder une marge de manoeuvre afin d'avoir suffisamment de temps pour bien préparer chaque étape.

Il est entendu que chacune des parties de la planification doit être en cohérence avec les autres. En fait, toute la planification doit être congruente avec la décision à prendre.

Une fois la planification faite, une révision de celle-ci n'est pas superflue pour s'assurer que rien ne cloche une fois la démarche amorcée.

3.6 Références

ALLAIRE, François, Henri, TARDIF et Marcel VEILLETTE. «La planification et l'évaluation: la rationalité comme objectif et outil de changement», in TESSIER, Roger et TELLIER, Yvan. *Changement planifié et développement des organisations. Théorie et pratique*, Montréal, Les Éditions de l'I.F.G., 1973, 825 p.

ANDERSON, Scarvia, et John S, HELMICK, Eds. *On Educational Testing*, San Francisco, Jossey-Bass Publ., 1983, 127 p.

BISQUERET, Claire, et José Luis, WOLFS. «Propositions de quelques procédures de notation. Comment noter nos élèves?», *Revue belge de Psychologie et de Pédagogie*, Tome 46, nos 185-186, 1984, p. 35-62.

Exemple d'un tableau de consignation pour un ensemble d'objectifs.

Élèves	Objectifs			Remarques
	Identifier les personnages	Identifier les caractéristiques merveilleuses	Identifier ce qui évoque des moeurs	
Julie	X	X	X	
Stéphanie	X		X	
Benoît	X	X	X	
Pierre		X	X	
...				

X: Objectif atteint

BLOOM, Benjamin S., J.-Thomas, HASTINGS, et George F., MADAUS. *Handbook on Formative and Summative Evaluation of Student Learning*, New-York, McGraw-Hill, 1971, 923 p.

CARDINET, Jean. *Pour apprécier le travail des élèves*, Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1984, 75 p.

CARDINET, Jean. *Des instruments d'évaluation pour chaque fonction.* Neuchâtel, Institut romand de recherches et de documentation pédagogiques, 1983, 81 p.

CARDINET, Jean. «L'adaptation des tests aux finalités de l'évaluation», *Sciences de l'éducation*, 2 – 3, 1973, p. 148-182.

DE KETELE, Jean-Marie. «Le passage de l'enseignement supérieur: les facteurs de réussite», *Humanités chrétiennes*, n° 4, 1983, p. 294-306.

DE KETELE, Jean-Marie. *Quelles questions dois-je me poser lorsque je construis une situation d'évaluation*?, Document 6, ICAFOC-CRAF, 03-10-1983.

GLASS, Gene V., et Julian C., STANLEY. *Statistical Methods in Education and Psychology*, Englewood Cliffs, Prentice-Hall, 1970, 596 p.

GOUVERNEMENT DU QUÉBEC, Conseil supérieur de l'Éducation. *L'évaluation: Situation actuelle et voies de développement.* Rapport 1982-1983 sur l'état et les besoins de l'éducation, Québec, 1983.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction de l'évaluation pédagogique. *«60% de quoi?» Réflexion sur la notation, la consignation et la transmission des résultats scolaires d'évaluation sommative, Classes du secondaire* 16-7508, Québec, janvier 1983, 49 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction de l'évaluation pédagogique. *Document de réflexion. Réflexion sur la notation, la consignation et la transmission des résultats scolaires d'évaluation formative et d'évaluation sommative, Classes du primaire*, Québec, septembre 1983, 71 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction de l'évaluation pédagogique. *Guide pour la construction d'instruments de mesure dans le cadre de la Banque d'instruments de mesure*, Chapitre II – III, Document de travail, 16-7300-02, B.I.M., Québec, septembre 1984, 38 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation, Direction du développement de l'évaluation. *Guide pour l'établissement du seuil de réussite*, 16-7514, Québec, mars 1986, 56 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale de l'évaluation, Direction du développement de l'évaluation. *Vers un bulletin scolaire renouvelé. Orientations du ministère en notation, consignation, transmission et conservation des résultats scolaires, Classes du primaire et du secondaire*, Québec, juillet 1985, 59 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale du développement pédagogique. *Éléments de docimologie: Lexique*, Québec, 1985, 23 p.

JAEGER, Richard M. *Measurement Consequences of Selected Standard-Setting Models*. Symposium on Measurement issues related to performance standards in competency —based education, National Council on Measurement in Education, San Francisco, avril, 1976, p. 22-27.

LAVERGNE, Jean-Paul. *La décision. Psychologie et méthodologie*, Paris, Les éditions ESF, 1983, 144 p.

LIVINGSTON, Samuel A. et Michael J., ZIERY, *Passing Scores, A Manual for Setting Standards of Performance on Educational and Occupational Tests*, Princeton, N.J., Educational Testing Service, 1982, 70 p.

LIVINGSTON, Samuel A. «Comments on Criterion-Referenced Testing», *Applied Psychological Measurement*, Vol. 4 N° 4, 1980, p. 575-581,

MILLMAN, Jason. «Passing Scores and Test Lenghs for Domain-Referenced Measures», *Review of Educational Research*, Vol. 43, N° 2, 1974, p. 205-216.

MESKAUSKAS, John A., NORCINI, John J. «Standard-Setting in Written and Interactive (oral) Specialty Certification Examinations», *Evaluation & the Health Professions*, Vol. 3, N° 3, September 1980, p. 321-360.

MORISETTE, Dominique. *La mesure et l'évaluation en enseignement*, Québec, Les Presses de l'université Laval, 1984, 423 p.

NITKO, Anthony J. *Educational Tests and Measurement an Introduction*, Toronto, Harcourt Brace Jovanovich, 1983, 661 p.

POPHAM. W. J. *Modern Educational Measurement*, Englewood Cliffs, N.J. Prentice-Hall, 1981, p. 371-399.

SHEPARD, Lorrie. «Standard Setting Issues and Methods», *Applied Psychological Measurement*, Vol. 4, N° 4, 1980, p. 447-467.

TOURNEUR, Yvan. «La certification des compétences», *Mesure et évaluation en éducation*, vol. 7, n° 4, 1985, p. 5-20.

TUCKMAN, Bruce W. *Measuring Educational Outcomes. Fundamentals of Testing*, Harcourt Brace Jovanovich, 1975, 523 p.

ZIEKY, Micheal J., et Samuel A., LIVINGSTON, *Basic Skills Assessment. Manual for Setting Standards on the Basic Skills Assessment Tests*, Princeton, New Jersey, Educational Testing Service, 1977.

La mesure en évaluation sommative

Préambule 52
4.1 Quelles qualités rechercher pour une épreuve? 52
A. La validité 52
B. La fidélité 54
4.2 Que faut-il surveiller au moment de la passation de l'épreuve? 55
4.3 Quelles précautions faut-il prendre lors de la correction des réponses? . . . 59
4.4 Comment interpréter les données? . . 60
A. L'analyse d'items 60
B. L'interprétation des données. . . 62
4.5 Quelles qualités assurer quand on prend les informations complémentaires? 64
4.6 Résumé 64
4.7 Références 65

Chapitre quatrième

Dans ce chapitre, nous verrons que:

1. Pour vérifier la performance des élèves on a recours à des épreuves de toutes sortes.
2. Pour rassembler les renseignements complémentaires on procède généralement par observation plus ou moins systématique.
3. Les résultats d'une épreuve doivent représenter avec exactitude la compétence de l'élève.
4. L'épreuve utilisée doit bien mesurer ce qu'elle prétend mesurer et rien de plus.
5. L'épreuve utilisée doit permettre d'obtenir des résultats constants.
6. L'épreuve utilisée doit favoriser une interprétation valable et être facile à utiliser.
7. On doit placer l'élève dans les meilleures conditions possibles pour passer l'épreuve.
8. De nombreux facteurs peuvent être sources d'erreurs dans la mesure.
9. L'analyse des items contribue à augmenter la validité des résultats d'une épreuve.
10. On interprète les résultats des élèves en rapport avec les objectifs vérifiés.
11. Dans tous les cas, on ne doit jamais utiliser aveuglément le seuil de réussite.
12. Lors de la prise d'informations complémentaires, on s'assure de leur pertinence, de leur objectivité et de la suffisance de leur nombre.

Préambule

Lors de la planification, on s'est préoccupé de bien préparer la réalisation de l'évaluation sommative. On a préparé le plan qui précise les décisions à prendre et les règles à suivre pour y arriver, qui indique les jugements qu'il faudra faire ainsi que les règles à suivre pour porter les jugements et enfin, qui désigne les informations à rassembler de même que les moyens pour les obtenir. Dans ce chapitre, nous voyons comment faire pour obtenir des informations de qualité en vue des jugements à porter. On met donc l'accent sur la qualité de la réalisation et sur ce qu'on fera si on rencontre des problèmes non prévus lors de la planification.

Souvent l'interprétation que l'on fait des résultats d'une épreuve ne fait que confirmer le jugement de compétence que l'on peut porter si l'on connaît bien l'élève. C'est la règle courante et la situation contraire doit susciter une vérification de notre part. Cette épreuve n'est toutefois pas inutile et mérite d'être préparée et utilisée avec soin car elle donne à l'élève une autre occasion de faire valoir ses acquisitions, occasion qui est souvent la dernière; elle est utile aussi à la société en lui permettant de contrôler les compétences transmises par l'école; enfin elle fournit des données plus objectives à l'enseignante ou à l'enseignant pour éclairer son jugement.

Pour des fins sommatives, il s'agit de faire le bilan des apprentissages prévus par le programme d'études. On est loin alors d'une certaine «moyenne» des réalisations de l'élève au cours de son apprentissage. On vise plutôt à rassembler diverses informations sur ce que l'élève peut faire en fin de course.

4.1 Quelles qualités rechercher pour une épreuve?

La principale qualité que doivent posséder les résultats d'une épreuve c'est leur capacité de représenter avec justesse la compétence de l'élève en vue de prendre les meilleures décisions.

A. La validité

On doit s'assurer que l'épreuve que l'on utilise mesure bien ce qu'elle prétend mesurer et rien de plus, c'est-à-dire que ses résultats soient valides et qu'il soit possible d'en faire une interprétation valide. On présente au tableau 4 des questions que l'on peut se poser pour vérifier la validité d'une épreuve ou d'un ensemble d'épreuves avant leur utilisation. On remarque dans ce tableau que l'on tient compte de plusieurs types de validité selon la décision à prendre.

D'autres facteurs influencent à notre insu la validité des résultats. Certains dépendent de l'élève même, attitudes et motivation par exemple, alors que d'autres lui sont extérieurs comme le temps accordé à la passation d'une épreuve. Prenons par exemple une épreuve de compréhension écrite en français où plusieurs facteurs sont susceptibles d'influencer les résultats.

- Biais technique: Si un élève connaît beaucoup plus que les autres le sujet du texte utilisé dans l'épreuve, il est probable qu'il le comprenne plus en profondeur et ait de meilleurs résultats.
- Vitesse: Si on n'a pas prévu suffisamment de temps, ce qu'on mesure en réalité ce n'est plus la compréhension de texte mais la vitesse de lecture .
- Attitudes et motivation: Si l'épreuve est mal présentée, il ne s'agit plus d'une mesure pure de la compréhension. Les élèves très motivés réussiront mieux que

Tableau 4

Liste de questions pour vérifier la validité d'une épreuve avant de l'utiliser.*

Aspect de la validité	Questions possibles	Suggestions pour trouver des réponses
Validité de contenu:	Est-ce que tous les items font partie du domaine?	On vérifie la définition du domaine de l'épreuve.
– représentativité du domaine	L'ensemble des items représente-t-il les caractéristiques essentielles du domaine (contenu et habileté intellectuelle) et l'ensemble des questions possibles?	On vérifie si chaque item (forme et contenu) est congruent avec le domaine.
	Y a-t-il suffisamment d'items pour que l'on puisse dire que la performance à l'examen représente la performance au domaine, aux apprentissages visés?	On vérifie si l'ensemble des items représente adéquatement le domaine.
	Les items (forme et contenu) et les clés de correction sont-ils jugés adéquats par les spécialistes?	
– pertinence par rapport à l'enseignement donné	Les items de cette épreuve ainsi que les habiletés intellectuelles ou psychomotrices qu'ils requièrent représentent-ils le domaine tel que nous l'avons défini (notre conception du programme, notre définition de l'intelligence, notre conception de ce trait de personnalité)?	On écrit la définition que l'on se donne de ce domaine. On vérifie chaque item et l'ensemble de l'épreuve comme précédemment mais cette fois-ci par rapport à sa propre définition du domaine.
Validité de construit	Est-ce qu'on peut interpréter les résultats comme étant significatifs par rapport à l'habileté ou au trait que l'on veut mesurer? Qu'est-ce que l'épreuve mesure vraiment?	On se réfère à la logique et à l'expérience pour expliquer les performances fortes ou faibles à l'épreuve (plausibilité de l'interprétation, relation ou non-relation prévue avec d'autres données, influence prévue ou non de certains facteurs). Il n'est pas nécessaire que toutes les explications soient cohérentes avec l'interprétation que l'on recherche. On procède à des études empiriques, expérimentales ou de corrélation, pour vérifier les explications faites à l'étape précédente. On utilise toutes les autres informations que l'on peut établir sur la validité pour interpréter les résultats.
Validité pour la sélection	Les résultats à l'épreuve peuvent-ils être utilisés pour prévoir la situation d'un élève par rapport à un autre domaine?	On analyse les tâches auxquelles on se réfère (travail, option, programme). On développe l'épreuve et on l'essaie sur un échantillon d'élèves. On vérifie la relation avec le référent (corrélation). On étudie les erreurs de classification qui sont faites. On analyse les habiletés qui doivent être développées ainsi que les exigences des différentes classes.
	Les résultats peuvent-ils être utilisés pour sélectionner des personnes pour un travail, un programme, etc.? Quelle est la grandeur de l'erreur?	
Validité de placement	L'épreuve peut-elle être utilisée pour placer des élèves dans différentes classes?	On développe une épreuve pour prédire les différents niveaux de succès dans les différents programmes. On vérifie si les placements que l'on fait à partir des résultats à l'épreuve sont confirmés.

* Adapté de Nitko, 1983, p. 438

ceux qui le sont moins, les premiers ne se laissant pas arrêter par cette mauvaise présentation.

- Habileté à passer des examens: Si des élèves ont développé des stratégies efficaces pour passer des examens, ils réussiront mieux que ceux qui sont aussi habiles en lecture mais qui n'ont pas développé de telles stratégies.
- Directives à l'élève: si un élève ne comprend pas bien les directives de l'examen, ses résultats pourront en être affectés.
- Habileté à indiquer ses réponses: Le type de feuille-réponses influence les résultats à l'épreuve (rapidité et justesse pour indiquer la bonne réponse).
- Seuil de réussite inadéquat.

On voit donc que bien des sources d'erreurs existent sur lesquelles on a peu de contrôle. Toutefois si on est perspicace on peut arriver à en contrôler quelques-unes et à diminuer ainsi la marge d'erreur.

B. La fidélité

Une autre qualité qui contribue à la validité d'une épreuve est la fidélité, c'est-à-dire la constance des résultats à cette épreuve. En effet, quelle confiance pourrions-nous avoir en une épreuve si les résultats des élèves variaient d'une fois à l'autre? On peut toujours accepter une légère variation mais ce qu'on vise, c'est de connaître la vraie performance de l'élève. Deux facteurs peuvent influencer la constance des résultats des élèves: le contenu ou l'échantillon spécifique des items utilisés et les circonstances de passation de l'épreuve.

Ces deux facteurs peuvent aussi se combiner. Par exemple, un élève peut réussir 8 et 10 questions à deux épreuves portant sur le même contenu mais étant de difficulté différente. Supposons encore qu'un élève passe deux fois la même épreuve. La première fois, il a une forte grippe et la deuxième fois, il est en pleine forme; ses résultats sont différents. Quel est le résultat qui représente vraiment son habileté?

Lorsque l'on choisit une épreuve, on se préoccupe des indices de fidélité qui sont rapportés dans sa description. Plusieurs méthodes sont utilisées pour estimer l'indice de fidélité. Le tableau 5 présente celles qui sont généralement utilisées lorsque l'on se situe dans une approche normative. Cette présentation comprend aussi une série de questions à se poser pour vérifier la fidélité.

Lorsqu'on se situe dans une approche critériée par ailleurs, il existe trois autres type de fidélité:

- la constance de l'identification des élèves compétents et non compétents à partir de l'application d'épreuves parallèles ou d'applications répétées d'une même épreuve (fidélité de décision);
- la constance des différences entre les scores des élèves et le seuil de performance obtenue à partir de deux épreuves parallèles ou aléatoirement parallèles (fidélité de variation);
- la constance des scores des élèves à deux épreuves parallèles (fidélité d'estimation).

On présente ces trois types d'indices de fidélité au tableau 6. On peut en trouver une description dans l'article de Serge B. Racine «La validité et la fidélité dans la mesure critériée» (1982).

Lorsque l'on choisit ou élabore une épreuve on se préoccupe aussi de la **qualité de l'interprétation** qu'elle permet de faire des données recueillies.

Voici une liste des questions qui permettent de vérifier des facteurs pouvant influencer la qualité de l'interprétation des résultats d'une épreuve:

- Comment les résultats sont-ils reliés à la performance visée?

- Quel est le seuil de réussite et comment a-t-il été établi?
- La structure de l'épreuve (dimensions et items) correspond-elle à ce qu'on veut mesurer?
- À quoi peut-on comparer les résultats(résultats passés, résultats de groupes)?

Une autre qualité que l'on recherche dans une épreuve est sa facilité d'utilisation. Voici quelques questions qui permettent de la vérifier:

- L'épreuve est-elle suffisamment courte pour ne pas causer de fatique, de stress...?
- L'épreuve est-elle utilisable en classe (temps disponible, organisation de la classe, matériel requis)?
- Les procédés d'administration sont-ils clairs, standardisés?
- L'épreuve est-elle compréhensible et significative pour l'élève?

En fait, il serait prudent d'expérimenter l'épreuve dans une autre école par exemple ou une autre classe, en ayant soin d'inclure dans l'échantillon, des élèves de différents dégrés d'habileté. On peut aussi utiliser une épreuve semblable mais qui contient des items différents. Cela permet d'identifier des sources de difficultés indésirables causées par des ambiguïtés dans les directives données à l'élève entre autres. Ces expérimentations permettent de vérifier le niveau d'exigence de l'épreuve.

4.2 Que faut-il surveiller au moment de la passation de l'épreuve?

Si l'on veut favoriser chez l'élève une performance qui représente bien sa compétence, on doit lui fournir les meilleurs conditions possibles pour passer l'épreuve. Ces conditions comprennent la présentation de l'épreuve, le moment et la durée, l'atmosphère et la préparation immédiate et éloignée de l'élève.

— La préparation matérielle de l'épreuve

- Une copie claire, bien dactylographiée, organisée de façon logique facilite la tâche à l'élève bien préparé et lui évite de faire des erreurs. De plus, cela lui donne l'impression que l'enseignante ou l'enseignant trouve importante la passation de cette épreuve. La passation de tout autre genre d'épreuve doit également être bien organisée, que ce soit un laboratoire ou une performance en ski de fond par exemple.
- Chaque élève doit recevoir une copie(ne pas écrire de questions au tableau, ne pas se contenter de lire les questions car cela nuit à ceux qui ont des problèmes d'audition ou de vision).

Tableau 5

Types d'indices de fidélité d'une épreuve selon l'approche normative

Méthode d'estimation	Questionnement	Facteurs de fidélité
Alternative de formes sans intervalle de temps	**I. Influence de différents échantillons de contenu** a. Les scores sont-ils affectés par l'échantillonnage de contenu différent lors d'une même occasion? b. Est-ce que deux épreuves soigneusement «jumelées» sont interchangeables (équivalentes, parallèles)?	L'échantillonnage du contenu
Spearman-Brown double longueur	a. Les scores à deux moitiés d'épreuve sont-ils affectés par l'échantillonnage de contenu lors d'une même occasion? b. Qu'est-ce qui est un estimé du coefficient de fidélité des deux formes de l'épreuve?	L'échantillonnage du contenu
Kider-Richardson Formule 20 ou 21 et Coefficient Alpha de Cronbach	a. Les scores sont-ils affectés par l'échantillonnage de contenu différent lors d'une même occasion? (On ne se préoccupe pas de l'équivalence des parties de l'épreuve) b. Quel est l'estimé de l'homogénéité de l'épreuve? c. Quelle est la consistance des réponses d'un item à l'autre?	L'échantillonnage du contenu et l'homogénéité
Test-retest	**II. Influence des occasions** a. Comment les scores à un même échantillon de contenu sont-ils affectés par la passation de l'épreuve à un autre occasion? b. Les scores à cette épreuve sont-ils stables dans le temps?	L'échantillonnage du temps ou de l'occasion
Alternative de formes avec intervalle de temps	a. Les scores sont-ils consistants quelles que soient la forme utilisée ou l'occasion de l'administration? b. Les scores sont-ils stables par rapport au temps et aux échantillons de contenu?	L'échantillonnage du temps ou de l'occasion et l'échantillonnage du contenu
Fidélité du correcteur	**III. Influence de différents correcteurs** a. Les scores seront-ils différents si on utilise différents correcteurs? b. L'épreuve est-elle objective? c. Les résultats attribués par différents correcteurs sont-ils interchangeables?	L'échantillonnage du correcteur

Inspiré de Nitko, 1983, p. 401.

Tableau 6

Les caractéristiques des scores et des décisions pour les trois types de fidélité selon l'approche critériée*

Éléments d'analyse	Types de fidélité		
	Fidélité de décision	**Fidélité de variation**	**Fidélité d'estimation**
Interprétation des scores	Le score individuel par rapport au score de césure	Le score individuel par rapport au score de césure	Le score individuel est utilisé pour estimer le score relié au domaine
Type de décision ou d'information requise pour la décision	L'identification des sujets compétents ou incompétents	Classement des sujets compétents ou incompétents le long d'un continuum	Le niveau de compétence relié au domaine
Risque associé aux erreurs de décision	Les risques sont également sérieux	Les risques reliés au mauvais classement des élèves qui sont loin au-dessus ou au-dessous du score de césure sont moins sérieux que pour les autres	La valeur des risques varie selon les méthodes d'estimation

* Inspiré de Racine, 1982, p. 97

- Bien sérier les questions selon leur type ou leur niveau de difficulté.
- Donner des directives claires à l'élève: nombre d'items, temps alloué...

— Le moment

Il faut s'efforcer de choisir un moment favorable pour la passation de l'épreuve, un moment où l'on sait que l'élève a le plus de chance d'être frais et dispos. Il est reconnu que les mardis, mercredis et jeudis matins sont propices à la passation d'une épreuve parce que les élèves sont mieux disposés qu'en après-midi ou que le vendredi.

— La durée

Comme il a été mentionné, l'élève doit disposer de suffisamment de temps pour compléter l'épreuve à moins qu'il s'agisse d'un test de vitesse. Il est toutefois très rare qu'on ait à avoir recours à des tests de vitesse en éducation. La plupart des situations de sanction exigent plutôt des épreuves de «puissance» c'est-à-dire qu'on cherche seulement à identifier ce que l'élève a acquis. Pour estimer la durée d'une épreuve, on se réfère au temps que les élèves les plus lents prennent en classe pour exécuter des tâches semblables. On peut aussi expérimenter l'épreuve auprès d'un autre groupe d'élèves.

— L'atmosphère

Il est entendu que la passation de l'examen ne doit pas être perturbée par

des situations d'indiscipline, des pannes d'électricité, etc. La température ne doit pas être étouffante ou glaciale. L'attitude du surveillant ou de la surveillante doit être sérieuse et refléter un certain cérémonial sans toutefois être trop stressante.

— La préparation immédiate de l'élève

En évaluation sommative on a besoin de connaître le maximum de ce que l'élève peut faire. On doit donc lui donner les informations suivantes:

- le moment de passation de l'épreuve;
- les conditions de passation;
- le contenu général sur lequel portera l'épreuve;
- l'importance de chaque partie du contenu;
- le type de performance exigé de la part de l'élève;
- la façon de corriger et de noter l'épreuve;
- l'importance de l'épreuve par rapport au jugement à porter et à la décision à prendre.

Ces informations permettront à l'élève de se préparer par rapport au contenu visé par l'épreuve.

— La préparation éloignée de l'élève

La préparation éloignée concerne le développement de l'habileté à passer une épreuve. Plus l'élève a développé cette habileté, plus il peut vraiment manifester sa compétence par rapport aux apprentisages visés par l'épreuve. Cette habileté générale est composée des habiletés spécifiques suivantes:

1. lire les directives ou être attentif à ce que dit la personne qui surveille;
2. décider combien de temps accorder à chaque item;
3. lire attentivement pour savoir rapidement ce qui est exigé pour la bonne réponse;
4. identifier les interprétations possibles pour une question et répondre selon l'interprétation qui est la plus problable;
5. sélectionner une réponse qui a du bon sens parmi les choix suggérés même si on connaît peu le contenu de l'item, répondre à une question ouverte même si on n'est pas certain de sa réponse;
6. décider d'attendre pour répondre à une question particulièrement difficile et y revenir par la suite (y faire une marque pour ne pas l'oublier);
7. bien retrouver la place des réponses sur une feuille-réponses séparée;
8. penser et organiser ses réponses avant de commencer à les écrire;
9. écrire de façon lisible;
10. relire sa copie et corriger les réponses mal écrites.

Il est utile à l'élève de connaître ces habiletés et d'essayer de les développer et il revient à l'enseignante ou à l'enseignant de le renseigner à ce sujet.

Toute ces informations aideront l'élève à contrôler l'anxiété qu'il peut ressentir face aux épreuves à passer. De plus, si on l'encourage à se consentrer sur la tâche à accomplir plutôt que sur les émotions négatives qu'il ressent (nervosité excessive, prévision de l'échec et de ses conséquences), il pourra avoir une performance qui reflète pleinement les apprentissages qu'il a developpés.

Dans tous les cas, à la fois l'enseignante ou l'enseignant et l'élève doivent savoir exactement dans quel but on fait passer un examen. C'est la transparence de chaque étape de l'évaluation qui en conditionne l'acceptation.

4.3 Quelles précautions faut-il prendre lors de la correction des réponses?

Une fois la passation de l'épreuve terminée, il s'agit de corriger chacune des réponses. On applique alors ce qui a été prévu lors de la planification. On compare la production de l'élève au modèle de référence que l'on s'est donné (clé de correction).

S'il s'agit d'une situation d'observation comme lors d'un exposé oral, d'une performance en éducation physique ou en musique ou encore de la réparation d'un moteur d'automobile par exemple, on utilise une grille d'observation. Ces situations d'observation demandent à l'observateur une série de comportements qui se déroulent assez rapidement. Ce sont les même qu'on effectue lors de l'utilisation d'une grille de correction pour questions ouvertes:

attention: porter son attention sur un objet, une action, un comportement;

description: définir, décrire pour lui-même ce qu'il perçoit;

comparaison: comparer la description qu'il s'est faite de la performance de l'élève à celle qu'il y a dans la grille d'observation;

appréciation: apprécier la performance;

consignation: consigner son appréciation selon ce qui est prévu dans le protocole d'observation .

Dans le cas des items à réponse choisie la correctrice ou le correcteur joue un rôle très simple: la réponse est très facilement interprétée comme bonne ou non. La correction est plus complexe lorsque l'on a décidé de tenir compte de la connaissance partielle et que les choix de réponses ont été pondérés. Toutefois, la tâche est allégée si l'on a recours à l'informatique. La situation est différente dans le cas des items à réponse construite et dans les situations d'observation où la personne qui corrige est beaucoup plus impliquée. Il se peut en effet qu'elle soit influencée par les informations qu'elle possède à propos de l'élève qu'elle observe ou de qui elle corrige la copie (effets d'assimilations). La correction peut aussi être influencée par les copies ou performances qui précèdent (effets de contraste).

Dans l'effet d'assimilation la personne qui corrige est influencée par les informations qu'elle possède et elle s'attend à retrouver chez l'élève un certain niveau de performance. Si cette attente est bien ancrée dans son esprit, elle percevra surtout ce qui confirme son attente. Par exemple si elle sait que tel élève ne fait pratiquement jamais de fautes, elle pourra ne pas voir du tout certaines erreurs faites par l'élève. Ses attentes l'amènent à avoir une perception sélective de la réalité. Les facteurs qui l'influencent sont entre autres le niveau social de l'élève, certaines caractéristiques personnelles de l'élève, les conditions dans lesquelles s'est déroulée l'épreuve. La correctrice ou le correcteur réduit donc de façon tout à fait inconsciente la distance qu'il y a entre ce que l'élève a produit réellement et ce qu'elle ou il attend de lui; c'est pourquoi on appelle cela l'effet d'assimilation.

L'effet de contraste est un tout autre processus où pour mieux percevoir la production de l'élève la correctrice ou le correcteur la distingue de ce qu'elle ou il vient de corriger. Par exemple, si on vient de lire une copie franchement médiocre, la copie qui suit pourra paraître bien meilleure qu'elle ne l'est en réalité. Ou encore on sera plus sévère dans la correction de la deuxième moitié des copies. Ceci se produit aussi lors d'observations de comportements. Il arrive aussi qu'à l'intérieur de la même copie la personne qui corrige compare des éléments et qu'elle force là encore la différence entre chacun. En fait elle a

de la difficulté à garder stable, constant, le cadre de référence qu'elle s'est donné. Le fascicule 5 suggère certaines techniques de correction qui tentent de contrôler ces sources d'erreurs comme par exemple de changer l'ordre des copies d'un numéro à l'autre lors de la correction.

Dans les items à réponse construite, c'est le jugement humain qui détermine l'exactitude des réponses. Qu'arriverait-il si la correctrice ou le correcteur évaluait deux fois les mêmes copies? Est-ce que cela donnerait les mêmes résultats? L'écriture de l'élève influence-elle la personne qui corrige? Cette personne est-elle fatiguée? L'important est d'être conscient qu'il existe plusieurs sources d'erreurs et par conséquent qu'il faut être prudent. L'utilisation d'une grille de correction précise favorise une correction de qualité.

Si l'on décide de classer les réponses selon les catégories suivantes, il faut que l'on ait décrit à l'avance les caractéristiques des performances que ces catégories représentent:

1. des réponses qui sont **inadéquates** par rapport à ce qui est attendu (insuccès);
2. des réponses **tout juste adéquates** (seuil);
3. des réponses **adéquates** qui présentent, en plus, des **caractéristiques souhaitables** indiquant une meilleure qualité (cible);
4. des réponses **supérieures** à celles du niveau précédent (optimum).

4.4 Comment interpréter les données?

Interpréter les données, c'est les rendre significatives. Pour tirer de l'information des questions d'une ou de plusieurs épreuves, il faut que ces questions soient de qualité satisfaisante. En effet, comment des questions ambiguës pourraient-elles renseigner sur les apprentissages des élèves? Comment pourrait-on dans un tel cas juger de leur compétence? Il s'avère donc important de commencer par analyser chacune des questions à la lumière des réponses des élèves. On procède alors à une analyse d'items.

A. L'analyse d'items

L'analyse d'items* consiste à vérifier le degré de difficulté de l'item et son pouvoir de discrimination ou sa sensibilité à l'apprentisage selon l'approche choisie(normative ou critériée) de même qu'à étudier la façon dont les élèves ont répondu. C'est aussi l'occasion de vérifier à nouveau la congruence entre les items et ce qu'on voulait mesurer.

- Le degré de difficulté est le pourcentage d'élèves qui ont réussi l'item. Il indique jusqu'à quel point les élèves ont eu de la difficulté à répondre. À l'occasion du calcul de l'indice de difficulté, on en profite pour analyser les réponses des élèves. Si on constate que la majorité donne une réponse différente de la réponse prévue, on doit prendre le temps de vérifier si la question était vraiment claire, on doit aussi se demander si la bonne réponse ne serait pas celle donnée par les élèves. D'un autre côté, l'analyse des erreurs faites par les élèves peut renseigner beaucoup sur l'état de leurs apprentissages; il faut toutefois que les items à réponse choisie comprennent des leurres qui puissent apporter de telles informations.
- Le pouvoir de discrimination indique jusquà quel point un item permet de différencier les élèves forts des élèves faibles tels que départagés par les résultats obtenus à l'ensemble des items d'une

* Le fascicule 5 des *Éléments de docimologie* explique en détail les diverses méthodes d'analyse d'items.

épreuve. Par exemple, selon l'approche normative classique, on penserait qu'un item manqué par les élèves qui ont les meilleurs résultats à l'épreuve et réussi par ceux qui ont les moins bons résultats, est un item inadéquat. Un tel phénomène dénote souvent une ambiguïté dans la question. La sensibilité à l'apprentissage nous informe sur la capacité qu'a un item d'indiquer s'il y a eu apprentissage ou d'indiquer les divers degrés de réussite d'une performance. Par exemple, selon l'approche critériée, un item manqué par les élèves qui ont suivi le cours et réussi par ceux qui ne l'ont pas suivi ne contribue pas à distinguer les élèves compétents des élèves incompétents par rapport aux apprentissages visés.

- L'étude de la façon dont les élèves ont répondu à tous les items peut révéler des phénomènes importants. Par exemple, on peut constater que la majorité des élèves n'a pas répondu à certains items parce que l'examen était peut-être trop long ou trop difficile. Il se pourrait d'un autre côté qu'une partie du contenu soit systématiquement échouée. Cet état de fait pourrait entraîner la décision de ne pas tenir compte de ces résultats dans l'évaluation sommative finale. Il serait pertinent alors d'organiser des activités correctives et par la suite de poser de nouvelles questions pour vérifier l'apprentissage des élèves.

La correction des items à réponse construite est sujette à la subjectivité du correcteur; c'est pourquoi il est conseillé de vérifier la constance des corrections. Pour ce faire, on procède à une deuxième correction. Idéalement, on corrigerait à nouveau toutes les copies ou l'on demanderait à quelqu'un d'autre de le faire. À défaut, on corrige à nouveau 20% des copies. Par la suite on compare les corrections. Plus les résultats sont semblables, plus la correction est fidèle. Une clé de correction précise contribue à augmenter la correspondance entre les corrections.

Si l'analyse d'items montre que certaines questions ne sont pas valables, la seule chose à faire est d'annuler ces questions dans le calcul des résultats. Il faut alors vérifier les conséquences d'une telle mesure sur la cohérence de la démarche d'évaluation. Ce serait le cas par exemple si une partie importante du contenu n'était pas vérifiée. Si on prévoit utiliser à nouveau cette épreuve, il faut remplacer les items qui ont été éliminés.

Il est important de procéder à une analyse d'items lorsque l'on utilise des épreuves qui n'ont pas été expérimentées. L'étude des réponses des élèves nous permet d'identifier les questions non valables et de se faire une idée sur la valeur globale de l'examen. Par la suite, on interprétera les résultats des élèves en tenant compte d'informations plus valides que si on avait conservé les renseignements erronés que donnaient les mauvaises questions. Avant d'interpréter les données, on doit aussi procéder à d'autres verifications. On doit entre autres s'assurer que

— personne n'a pris connaissance de l'épreuve pour en divulguer la teneur au préalable (sécurité);

— la façon d'élaborer l'examen a été adéquate (planification, compétence des rédactrices et des rédacteurs, congruence avec les objectifs, etc.);

— la constance des résultats à l'examen est suffisante (fidélité);

— le contenu de l'examen représente ce qui a été enseigné en classe (validité);

— la façon d'établir le seuil de réussite n'a pas de failles (manque de temps, juges incompétents, etc.);

— le seuil de réussite est adéquat;

— l'examen n'a pas de défauts techniques (directives inadéquates, qualité de l'imprimerie, etc.).

B. L'interprétation des données

On commence l'interprétation des données quand on déclare qu'une réponse à une question est bonne ou mauvaise, ou encore quand on nuance en disant que telle réponse est excellente ou en accordant dix sur dix. Cette appréciation est donnée en se référant à l'objectif d'apprentissage visé et aux exigences fixées pour celui-ci. Ces exigences ont été établies lors de la planification de l'enseignement et de l'évaluation. Par prudence, on révise ces exigences à la lumière des résultats des élèves et de sa propre connaissance du programme.

On tient compte ensuite de plusieurs informations, plusieurs items, pour chaque objectif. Généralement lors de la préparation de l'épreuve , on prévoit les interprétations qu'on fera des résultats et on organise si possible les items de l'épreuve en conséquence. Il est certain que plus on a d'informations pertinentes, plus on est sûr de l'interprétation à donner et plus l'élève a l'occasion de manifester son niveau réel de performance .

Comme on se situe généralement en mesure critériée, les items de l'épreuve font référence à des objectifs en rapport à une performance attendue. Dans ce cas, la référence à un seuil de réussite est fort utile. Ce seuil peut être établi pour chaque objectif en rapport avec l'ensemble des items se rattachant à celui-ci ou pour l'ensemble de l'épreuve. Il peut être constitué d'un nombre d'items à réussir. Lors de la planification, on a choisi une méthode pour fixer le seuil de réussite. On pourrait avoir choisi une méthode a priori. Dans ce cas, le seuil fixé avant la passation de l'épreuve doit toujours être confronté aux conséquences qu'il entraîne, nombre exagéré d'échecs par exemple. Si par contre on a opté pour une méthode où le seuil de réussite est fixé à partir des résultats à l'épreuve (méthode a posteriori), il s'agit d'appliquer cette méthode selon la valeur du seuil trouvé.

Dans tous les cas, on ne doit jamais utiliser aveuglément le seuil de réussite. Plusieurs précautions sont à prendre. On doit:

- tenir compte de la marge d'incertitude autour du score choisi comme seuil de réussite;
- revoir le seuil de réussite si on a l'impression que trop d'élèves compétents n'atteignent pas ce seuil ou que trop d'élèves incompétents l'atteignent;
- réviser le seuil de réussite s'il a été fixé avant l'élimination de certains items.

Une fois le seuil établi, on procède à l'interprétation définitive des données de l'épreuve.

On compare la performance de l'élève à ce seuil et on consigne l'interprétation obtenue. L'information est encore plus riche si on interprète les performances de l'élève selon différents degrés de qualité par exemple, selon les quatre catégories suivantes:

— performance supérieure;
— performance souhaitée, désirée;
— performance minimale;
— performance inadéquate.

Tout dépend des jugements que l'on aura à porter par la suite. Si l'on décide de procéder de la sorte, il faut alors fixer les seuils qui correspondent à ces catécories.

L'utilisation de tableaux de résultats facilite l'application des seuils. On en présente un exemple à la page suivante. On indique pour chaque élève, le nombre d'items réussis en regard de chaque objectif . L'exigence établie pour réussir chaque objectif y figure aussi. Les objectifs 6 et 7 ont été vérifiés par des productions personnelles des élèves qui ont été notées à l'aide d'une échelle alphabétique.

On présente dans le deuxième exemple une façon de comparer le nombre d'objectifs atteints par les élèves. On y lit que 9% des élèves ont atteint seulement deux objectifs et que 93% en ont au moins atteint deux.

Exemple de tableau facilitant l'interprétation des données d'une épreuve

	Objectifs						
	1	2	3	4	5	6	7
Seuils de réussite: Nombre exigé d'items réussis	5/5	8/10	6/10	3/3	7/10	B	C
Élèves	Nombre d'items réussis ou cote obtenue						
Éric	4	9*	10*	3*	5	A*	A*
Isabelle	5*	8*	6*	3*	7*	A*	B*
Marie-Josée	5*	8*	6*	3*	7*	B*	B*
Martin	5*	6	7*	3*	4	B*	C*
Patrice	3	8*	6*	3*	3	C	C*

* = réussite de l'objectif

Exemple d'approche normative – pourcentages cumulés

Nombre d'objectifs atteints	Pourcentage d'élèves ayant atteint chaque nombre d'objectifs	Pourcentage cumulé
0	1%	100%
1	6%	99%
2	9%	93%
3	55%	84%
4	20%	29%
5	9%	9%

4.5 Quelles qualités assurer quand on prend les informations complémentaire?

Il a déjà été mentionné que pour juger de la compétence de l'élève, il faut tenir compte de toutes les informations pertinentes. Ces renseignements portent sur des événements précis, le déroulement des apprentissages, la valeur des épreuves utilisées, les circonstances de passation des épreuves ou certaines caractéristiques des élèves.

Certaines précautions sont à prendre dans le choix et la prise de ces renseignements. On doit avant tout s'assurer de leur **pertinence** par rapport aux jugements à porter et aux décisions à prendre particulièrement en ce qui concerne les caractéristiques de l'élève. En effet, dans ce domaine il faut toujours avoir validé, lors de la planification, les choix faits, c'est-à-dire les informations à rechercher et les éléments à observer. Il ne faut pas oublier que l'observation, des intérêts et des attitudes entre autres, comporte une part de subjectivité et dépend de la définition que l'on se donne de ce qu'on veut observer. Pour favoriser l'**objectivité** de ces prises d'informations on peut :

- se faire une liste des éléments à observer et s'y référer régulièrement;
- noter au fur et à mesure ses observations et ne pas se fier seulement à sa mémoire;
- comparer ses observations à celles d'une autre observatrice ou d'un autre observateur quand c'est possible.

Il est prudent d'effectuer les observations en **quantité suffisante** pour assurer la représentativité des renseignements que l'on prend et la valeur des interprétations qu'on en fera.

4.6 Résumé

Dans ce chapitre nous avons vu comment procéder aux différentes prises d'informations et quelles précautions prendre. Il est évident que la prise de l'information doit se dérouler dans les conditions les plus favorables possible. Lors de l'utilisation d'épreuves, il s'agit de bien préparer les élèves et de leur fournir une épreuve de qualité. En deuxième lieu on s'assure d'une correction valide, objective et fidèle en utilisant une clé de correction précise ou une grille d'observation de même que des façons de faire qui favorisent l'objectivité et la constance.

Pour l'interprétation des données, on tire avantage d'une bonne analyse et d'une organisation éloquente des résultats de l'élève. Des listes d'objectifs ou des tableaux de résultats aident à visualiser globalement les données par rapport aux dimensions retenues pour l'évaluation. On compare alors les résultats des élèves aux exigences fixées en rapport aux jugements à porter et aux décisions à prendre.

Lors de la prise d'informations complémentaires, il faut assurer la pertinence des informations que l'on prend, l'objectivité de l'observation et la représentativité des événements pris en compte. Pour ce faire, on systématise les prises d'informations et l'on compare ses observations à celles d'autres observateurs quand cela est possible.

Dans tous les cas, il est recommandé d'être prudent compte tenu des nombreuses sources d'erreurs qui influencent les processus de prises d'informations. Cette prudence sera une garantie pour la justice des jugements qui suivront.

4.7 Références

ALLAIRE, François, Henri TARDIF et Marcel, VEILLETTE. «La planification et l'évaluation: la rationalité comme objectif et outil de changement», in TESSIER, Roger et TELLIER, Yvan. *Changement planifié et développement des organisations. Théorie et pratique,* Montréal, I.F.G, 1973, 825 p.

ANDERSON, Scarcia, et John S., HELMICK. Eds. *On Educational testing,* San Francisco, Jossey-Bass Publ. 1983, 127 p.

BISQUERET, Claire, et José Luis, WOLFS. «Proposition de quelques procédures de notation. Comment noter nos élèves?», *Revue belge de Psychologie et de pédagogie,* Tome 46, nos 185 -186, 1984, p. 35-62.

BLOOM, Benjamin S., J.Thomas, HASTINGS, et George F., MADAUS. *Handbook on Formative and Summative Evaluation of Student Learning,* New York, McGraw-Hill, 1971, 923 p.

BRENNAN, Robert L., et Micheal T., KANE. «Signal/Noise Ratios for Domain-Reference Tests», *Psychometrica,* Vol. 42 , N° 4 , décembre 1977, p. 609-625.

CARDINET, Jean . *Pour apprécier le travail des élèves,* Neuchâtel, Institut romand de recherches et de documentation pédagogiques , 1984, 75 p.

CARDINET, Jean . *Des instruments d'évaluation pour chaque fonction,* Neuchâtel, Institut romand de recherches et de documentation pédagogique, 1983, 81 p.

GOUVERNEMENT DU QUÉBEC, Conseil supérieur de l'Éducation . *L'évaluation : situation actuelle et voies de développement.* Rapport 1982-1983 sur l'état et les besoins de l'éducation, Québec, 1983.

DE KETELE, Jean-Marie. «Le passage de l'enseignement secondaire à l'enseignement supérieur: les facteurs de réussite», *Humanités chrétiennes,* n° 4, 1983, p. 294-306.

DE KETELE, Jean-Marie. *Quelles questions dois-je me poser lorsque je construis une situation d'évaluation?* Document 6, ICAFOC -CRAF , 03-10-1983.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction de l'évaluation pédagogique. *60% de quoi ? Réflexion sur la notation, la consignation et la transmission des résultats scolaires d'évaluation sommative, Classes du secondaire,* Document 16-7508, janvier 1983, 49 p.

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction de l'évaluation pédagogique. *Document de réflexion. Réflexion sur la notation, la consignation et la transmission des résultats scolaires d'évaluation formative et d'évaluation sommative, Classes du primaire,* septembre 1983, 71 p.

KESSNER, Christine. *L'évaluation du savoir écouter,* Neuchâtel, Institut romand de recherches et de documentation pédagogiques, mai 1983, 9 p.

LAVERGNE, Jean-Paul. *La décision. Psychologie et méthodologie,* Paris, ESF, 1983, 144 p.

LIVINGSTON, Samuel A. *Basic Skills Assessment. Manual for Setting Standards on the Basic Skills Assessment Tests,* Princeton, Educational Testing Service, 1977.

LIVINGSTON, Samuel A. et Micheal J., ZIEKY. *Passing scores. A Manual for Setting Standards of Performance on Educational and Occupational Tests,* Princeton, N.J., Educational Testing Service, 1982, 70 p.

LIVINGSTON, Samual A. «Comments on Criterion-Referenced Testing», in *Applied Psychological Measurement* ,Vol. 4, Nº 4 , 1980, p. 575-581.

MILLMAN, Jason. Passing Scores and Test Lenghts for Domain-Referenced Measures , *Review of Educational Research,* Vol. 43, Nº 2, 1974, p. 205-216.

MESKAUSKAS, John A., et John J., NORCINI. «Standard-Setting in Written and Interactive (oral) Specialty Certificate Examinations», *Evaluation & The Health Professions,* Vol. 3, Nº 3, September 1980, p. 321-360.

NITKO, Anthony J. *Educational Tests and Measurement. An Introduction,* Toronto, Harcourt Brace Jovano-vich,1983, 661 p.

NOIZET, Georges et Jean Paul CAVERNI. *Psychologie de l'évaluation scolaire,* Paris, P.U.F.,1978, 231 p.

POPHAM, W.J. «Performance Standards», *Modern Educational Measurement,* 1981, p. 371-399.

RACINE, Serge B. «La validité et la fidélité dans la mesure critériée», *Mesure et évaluation en éducation,* vol. 5 , nº 3, 1982, p. 92-110.

SHEPARD, Lorrie. «Standard Setting Issues and Methods», *Applied Psychological Measurement,* Vol. 4 Nº 4, 1980, p. 447-467.

THURLER, Monica. *Réflexion sur l'évaluation de l'élève,* Rapport d'atelier 1, Groupe de travail 2, Genève, Directeurs du projet SIPRI, févier 1983, 48 p.

TOURNEUR, Yvan. «La certification des compétences», *Mesure et évaluation en éducation,* vol. 7, nº 4, 1985, p.5-20.

TUCKMAN, Bruce W. *Measuring Educational Outcomes Fundamentals of Testing,* Hardcourt Brace Jovanovich, 1975, 523 p.

WOLFS, José Luis. *Quelques propositions de notation et d'évaluation certificative dans le cadre d'une relation éducative basée sur la transparence,* Rencontre Belgo-Franco-Suisse sur l'évaluation, Genève, 3-5 septembre 1984, 15 p.

Le jugement et la décision

Préambule 68
5.1 À quel moment juger de la situation de l'élève? 69
5.2 Comment juger? 69
A. Combiner les multiples données recueillies 69
B. L'erreur de jugement 71
C. Cas limites et cas imprévus . . . 72
D. La répartition des responsabilités . . 72
5.3 Comment prendre la décision? . . . 73
5.4 Comment consigner les jugements et les décisions? 74
5.5 Résumé 75
5.6 Références 76

Chapitre cinquième

Dans ce chapitre, nous verrons que:

1. À l'étape jugement, il s'agit d'intégrer les diverses informations recueillies et de statuer sur la compétence de l'élève ou sur d'autres aspects selon les règles de jugement établies lors de la planification.
2. Lors du jugement on doit suivre un processus rationnel, juste et valide.
3. Le jugement final se porte le plus près possible du moment où la décision sera prise.
4. Les jugements temporaires ont lieu le plus près possible des bulletins.
5. Lorsque l'on juge la situation de l'élève, on tient compte des interprétations faites à partir des épreuves, de la valeur même de ces épreuves, du vécu des élèves, de la valeur des seuils de réussite et des autres informations pertinentes.
6. On essaie de trouver le jugement qui sera le plus juste pour tel élève dans telles circonstances selon la décision à prendre.
7. L'enseignante ou l'enseignant ou le comité multidisciplinaire recommande à la direction de l'école la décision considérée la meilleure pour l'élève.
8. Le bulletin est un instrument de communication, on doit y utiliser un langage clair.

Préambule

Plusieurs choses ont été établies lors de la planification du jugement et de la décision. Dans ce chapitre, on attire l'attention sur la qualité de la réalisation de ces étapes. On décrit aussi de façon globale comment solutionner les problèmes imprévus qui peuvent surgir.

À l'étape mesure on s'est pourvu des informations nécessaires aux jugements à porter selon les décisions à prendre et on a en main différents renseignements sur le déroulement des apprentissages en cours d'année. Par exemple, on a comparé les performances de l'élève aux examens par rapport aux objectifs poursuivis, au domaine des tâches visées, à une performance standard ou à la performance des autres élèves. Ou encore, on a essayé de bien identifier son niveau de motivation, ses intérêts. Il s'agit maintenant d'intégrer ces différentes informations et d'arriver à statuer sur l'état de l'apprentissage de l'élève ou sur ses capacités.

Pour la sanction des études et la reconnaissance des acquis expérientiels, on juge si l'élève satisfait les exigences établies par le ministère de l'Éducation. Pour l'orientation dans la carrière scolaire et le passage à la classe supérieure, on juge aussi de l'état de l'apprentissage de l'élève et des autres aspects qui peuvent contribuer à sa réussite comme l'intérêt, le style et le rythme d'apprentissage, la motivation, la personnalité.

On applique alors aux moments fixés lors de la planification les règles de jugement qu'on y a établies. Par exemple à la fin de l'année, on examine les résultats obtenus par l'élève en rapport aux objectifs importants du programme d'études. On prend aussi en compte la valeur des examens et d'autres informations pertinentes. Par la suite, on juge de la compétence de l'élève en regard du programme d'études.

Quelle que soit la décision à prendre, il importe que le processus suivi lors du jugement soit rationnel, juste et valide.

5.1 À quel moment juger de la situation de l'élève?

Lors de la planification, en tenant compte de la nature des programmes d'études, on a fixé des moments où l'on portera les jugements temporaires et le jugement final pour chacun des programmes. Il s'agit alors d'appliquer cet échéancier en veillant à poser le jugement final le plus près possible du moment où la décision sera prise. Le jugement final se doit de par sa nature même, de refléter la compétence de l'élève à la fin du cours même s'il arrive souvent en milieu d'année qu'on ait à porter des jugements hâtifs pour permettre l'organisation de l'année scolaire suivante.

Si on le peut, on organise l'enseignement de sorte que l'élève garde actifs dans son répertoire les apprentissages importants qu'il fait tout au cours de l'année. On réserve pendant le déroulement de l'enseignement-apprentissage des moments où l'on place l'élève dans des situations où il doit utiliser des connaissances acquises et des habiletés développées. Les jugements temporaires pour leur part, doivent être portés le plus près possible de l'émission des bulletins en cours d'année, toujours dans le but de donner le portrait le plus récent de l'élève. En fait, tant que la période prévue pour l'apprentissage n'est pas terminée, l'élève est toujours susceptible de combler ses lacunes et d'atteindre un niveau satisfaisant. Les donnés recueillies en cours d'année ne sont pas cumulatives. L'humain étant un être changeant, la vie de l'élève devrait toujours être prioritaire à la facilité de la vie administrative.

5.2 Comment juger?

A. Combiner les multiples données recueillies

Lors de la planification, on a établi des règles de jugement. On a précisé l'importance que l'on accordait à chacun des éléments choisis et l'on a prévu plusieurs cas possibles. Maintenant, il s'agit d'appliquer judicieusement ces règles.

Quelle que soit la décision à prendre, orientation de l'élève dans sa carrière scolaire ou reconnaissance de ses acquis, on a toujours à combiner les différentes informations recueillies pour pouvoir estimer la situation de l'élève.

On a en main les interprétations que l'on a faites à partir des épreuves que l'on a utilisées. Pour porter le jugement, on prend alors en considération d'autres informations pertinentes:

- On tient compte de la valeur des épreuves. Si une épreuve s'avère invalide en elle-même ou pour un élève en particulier (apport socio-culturel, handicap physique, etc.), on ne peut pas tenir compte des résultats obtenus par les élèves à cette épreuve pour juger de leur compétence. Dans un tel cas, on se doit d'aller chercher les informations qui manquent par rapport aux apprentissages visés dans cette épreuve avant de porter le jugement. Ceci dans le but d'avoir toute l'information prévue dans le tableau-synthèse ou le tableau de spécifications.
- Si on s'est basé sur des seuils de réussite pour interpréter les résultats des élèves, à chacun des instruments de mesure, on doit vérifier la valeur de ces seuils.
- On consulte les résultats d'évaluation formative. Ceux-ci renseignent sur le déroulement de l'enseignement-apprentissage et permettent de mieux connaître l'élève.
- On se réfère au vécu de l'élève. Ce vécu a-t-il empêché l'élève de fournir lors de l'épreuve une performance représentative de sa compétence (maladie lors de la passation de l'épreuve, stress, etc.)? Les circonstances ont-elles nui à l'apprentissage de l'élève (interaction enseignant/

élève, troubles familiaux, etc.? On vérifie si certaines caractéristiques de l'élève n'ont pas nui à ses performances.

— Lorsque l'on tient compte d'informations sur l'intérêt de l'élève, sa motivation, ou d'autres aspects affectifs, on s'efforce de baser ses jugements sur des faits observés en nombre suffisant.

Il est important de recourir systématiquement aux règles de jugement préétablies lors de la planification, règles où les critères congruents avec la décision à prendre sont identifiés. Il est aussi très important de faire connaître ces règles afin de favoriser la justice.

Dans des situations très exceptionnelles, il est possible qu'on se voit dans l'obligation de ne pas appliquer les règles de jugement ou qu'on se doive de les examiner minutieusement. C'est le cas par exemple quand cela mène à un nombre inacceptable d'échecs ou à des jugements opposés aux jugements auxquels le côtoiement journalier avec l'élève nous a amenés. Il est possible en effet que des règles ne soient pas adéquates même si elles ont été établies avec rigueur et bonne volonté. Comme on ne peut être certain de la valeur de ces règles qu'après les avoir expérimentées, peut-être même plusieurs fois, il est prudent d'examiner les jugements auxquels elles donnent lieu.

Certaines règles de jugement peuvent être plus difficiles à appliquer que d'autres. La pratique d'additionner de nombreuses notes, d'en faire la moyenne et d'inscrire cette dernière au bulletin a longtemps été de mise. Le jugement proprement dit y avait peu de place. Même à cela, bien des enseignants trouvaient moyen d'adapter cette moyenne au jugement qu'ils faisaient quand la divergence était trop grande. Il demeure que dans cette pratique, l'étape «jugement» n'avait pas sa place attitrée et plusieurs ne faisaient que transcrire le résultat d'un calcul dans le bulletin. Actuellement, on est peu habitué à articuler de façon formelle et publique le jugement, ni à appliquer systématiquement des règles établies, et l'implication personnelle que cette pratique requiert fait ressentir plus intensément la responsabilité que l'on a dans les décisions qui en découleront. Ce qu'il ne faut pas oublier, c'est qu'il ne s'agit pas d'additionner les données recueillies mais d'arriver à les combiner pour enfin porter le jugement sur la compétence de l'élève en tenant compte des données pertinentes.

Lors de la planification des règles de jugement dans une approche critériée, on a décidé si on exigerait que l'élève soit compétent par rapport à tous les objectifs importants du programme, à la majorité des objectifs importants ou à une liste précise d'objectifs importants. Il s'agit de juger de la compétence de l'élève par rapport aux objectifs importants à partir des interprétations faites des résultats aux diverses épreuves. L'exemple suivant illustre la situation où l'on exige que l'élève soit compétent par rapport à la majorité des objectifs importants du programme d'études.

Si on applique cette méthode, on constate que seulement Sylvain n'a pas atteint la majorité des objectifs visés.

On pourrait aussi se référer au profil des apprentissages faits par l'élève. Dans ce cas, on exige que l'élève ait atteint certains objectifs jugés essentiels. L'utilisation de l'approche par profil n'est justifiée que lorsque l'on est certain de la nécessité de ces apprentissages. Si on utilisait cette méthode en exigeant que les objectifs 1 et 3 soient atteints, seulement Claude serait jugé compétent. Une telle règle de jugement est assez sévère. Elle doit donc être basée sur une analyse scientifique du programme et de la décision à prendre.

À chaque fois que l'on porte un jugement, il faut être conscient que les données recueillies à l'aide d'épreuves sont toujours

Exemple de résultats d'élèves

Élèves	Objectifs poursuivis 1	2	3	4	5	Jugements de compétence
Claude	R	R	R	R	E	oui
Mireille	R	R	E	R	E	oui
Monique	E	R	R	E	R	oui
Olivier	E	R	E	R	R	oui
Sylvain	E	E	E	E	E	non

R = Objectif réussi
E = Objectif non réussi

susceptibles d'être biaisées par plusieurs sources d'erreurs. C'est pourquoi il ne faut pas faire une simple transcription de ces données sans porter un jugement. Par ailleurs, ce jugement peut à son tour être influencé par d'autres sources d'erreurs qui biaisent aussi la perception et le jugement de l'évaluatrice ou de l'évaluateur. Cette personne doit s'efforcer de bien saisir la réalité en se centrant sur la décision pour laquelle elle porte le jugement. Quoique ne pouvant jamais se départir totalement de sa subjectivité, elle doit tenter de porter un jugement éclairé empreint de justice.

B. L'erreur de jugement

Un jugement erroné conduit à une décision tout aussi erronée, décision qui a des conséquences néfastes pour l'élève. Par exemple, la décision de placer un élève dans une classe où il ne peut pas suivre ou bien dans une classe où il s'ennuie parce qu'il n'y apprend rien.

Lors de l'application des règles de jugement il faut tenir compte des possibilités d'erreurs. Il s'agit de minimiser le plus possible les deux types de jugements erronés que l'on peut faire (voir le tableau 7), soit: juger qu'un élève n'est pas compétent par rapport aux apprentissages visés quand en réalité il l'est, ou juger qu'il est compétent quand en fait il ne l'est pas.

Il est malheureusement impossible d'éviter complètement les erreurs. Voici une liste de **précautions à prendre pour minimiser les erreurs de jugement**.

— Veiller à ce que le choix des exigences varie le moins possible selon les personnes qui les établissent. (Les opinions divergent. Il peut être dans l'intérêt de certains de maintenir des exigences élevées ou faibles.)

— S'assurer que l'ensemble des exigences représente bien le programme d'études.

— Baser ses jugements sur des informations représentant bien les apprentissages ou les caractéristiques à juger. Il se peut en effet que les informations ne correspondent pas bien au domaine ou aux caractéristiques particulières de l'élève.

— Baser ses jugements sur des informations assez nombreuses. Si on ne se base que sur un seul examen par exemple, il se peut qu'à une autre occasion l'élève ait un autre résultat.

— Tenir compte de toutes les informations pertinentes (maladie de l'élève, décès dans la famille

— S'assurer que les personnes qui jugent sont habiles à le faire et qu'elles le font de façon impartiale.

Tableau 7
Types de jugements erronés

Jugement porté	État réel de l'élève	
	L'élève a réalisé les apprentissages	L'élève n'a pas réalisé les apprentissages
Je juge que l'élève n'est pas compétent	Mauvais jugement: Erreur de type I (faux-négatif)	Bon jugement
Je juge que l'élève est compétent	Bon jugement	Mauvais jugement: Erreur de type II (faux-positif)

— Définir avec suffisamment de précision en quoi consiste la compétence de l'élève dont il faut juger pour tel ou tel programme.
— Demander aux personnes qui portent les jugements d'expliquer les principaux facteurs qui les guident.
— Travailler en équipe.
— Réviser annuellement les règles de jugement.
— Prévoir une façon de contrôler la qualité des informations sur l'élève (examens, conditions de passation,...).

C. Cas limites et cas imprévus

Comme nous l'avons vu pour certains élèves dans les exemples précédents, il arrive que la situation d'un élève ne soit pas facile à juger. C'est souvent le cas lorsque ses performances ont un niveau de qualité qui est tout juste acceptable et ceci après plusieurs prises d'informations. Jugera-t-on qu'il est suffisamment compétent ou non? Des circonstances imprévisibles peuvent aussi causer des problèmes. Que fait-on si un élève a un accident à la fin de l'année et ne peut passer les épreuves-synthèse prévues? Que fait-on si un élève passe les examens dans un état psychologique qui rend ses performances non représentatives? Que fait-on quand on sait que les interrelations entre un élève et la personne qui lui enseigne ont été si mauvaises que les données recueillies en cours d'année représentent plus la qualité de leur relation que les apprentissages réels de l'élève?

Dans tous ces cas difficiles à résoudre le souci de justice doit guider la personne qui juge. On essaie de trouver le jugement qui sera le plus juste pour tel élève dans telles circonstances pour telle décision à prendre. Dans ces cas, on a souvent plusieurs jugements à porter. Selon la décision à prendre, on jugera par exemple de la compétence de l'élève par rapport au programme d'études, de ses intérêts, de l'importance des apprentissages maîtrisés et de ceux qui ne sont pas maîtrisés pour la poursuite des études, des données recueillies en cours d'année, de la valeur des instruments de mesure utilisés, etc.

D. La répartition des responsabilités

Le partage des responsabilités a été établi lors de la planification en tenant compte des instructions émises par le ministère de l'Éducation et la commission scolaire. Au secondaire, le ministère de l'Éducation a une plus grande importance par les règles qu'il spécifie.

Il demeure qu'au primaire et même au secondaire, à l'intérieur de la marge de manoeuvre qui reste à l'école, ce sont souvent les cas spéciaux qui font ressortir le partage des responsabilités des jugements en évaluation sommative. De façon indirecte, les personnes qui ont élaboré les instruments de mesure et celles qui ont corrigé les réponses des élèves ont une responsabilité car elles contribuent à fournir des données qui serviront de base aux jugements.

Lorsque la situation est claire, c'est généralement l'enseignante ou l'enseignant qui juge de la compétence de l'élève par rapport aux apprentissages visés par le cours. Par ailleurs dans les cas douteux ou imprévus, habituellement on convoque un comité dont les membres analysent les données en main, confrontent leurs jugements en les justifiant. Ces comités peuvent comprendre un ou plusieurs enseignants, une conseillère ou différents professionnels, un conseiller en orientation, etc. Ce sera souvent le même comité qui prendra la décision.

Il est entendu que l'on doit fournir à l'enseignante ou à l'enseignant le soutien nécessaire à la tâche de porter ces jugements. Certains peuvent avoir besoin de formation par rapport au programme d'études concerné ou par rapport à la démarche d'évaluation sommative.

L'enseignante ou l'enseignant ou, selon le cas, le comité multidisciplinaire, devrait toujours vérifier si les exigences fixées lors de la planification sont valables (justes, réalistes, applicables, etc.). Si, après une étude approfondie, il s'avère nécessaire de modifier les exigences, on doit le faire.

On constate que fréquemment l'enseignant ou l'enseignante ne participe pas à un bon nombre de décisions prises à la suite des jugements qui ont été faits. C'est ainsi que dans bien des cas, l'orientation se précise plusieurs mois après la fin d'un cours; parfois une matière préalable n'est utilisée que plusieurs années après son étude, etc. Cependant, l'enseignante ou l'enseignant conserve toujours une grande responsabilité sur ces décisions parce qu'elles prennent leur source, en tout ou en partie, dans le jugement qu'elle ou qu'il a déjà porté sur la réalisation des apprentissages associés à tel ou tel programme.

5.3 Comment prendre la décision?

Dans la section traitant de la planification de l'étape décision (3.1), nous avons vu que les décisions qui relèvent d'une évaluation sommative sont très nombreuses:

— donner ou refuser un diplôme;
— placer l'élève dans un programme particulier: douance, récupération, année de transition, école alternative, etc.;
— faire passer l'élève à la classe supérieure, lui faire sauter une année, lui faire recommencer un cours, etc.;
— faire suivre des cours d'été;
— accorder les crédits rattachés au programme d'études;
— reconnaître les acquis expérientiels.

Dans la plupart des cas toutefois, la décision est simple à prendre: l'élève maîtrise tous les apprentissages essentiels ou il n'en maîtrise presque pas. Les règles de décision s'appliquent de façon quasi automatique.

On a identifié deux types d'erreur de décision. À chaque fois que la décision est douteuse, on peut se demander quelles sont les conséquences de chaque type d'erreur possible. On se demande, par exemple, quelles seraient les conséquences pour telle ou tel élève si on décidait de lui faire recommencer telle année, si en fait les apprentissages requis pour réussir la classe supérieure sont suffisamment maîtrisés, ou si on décidait de le faire passer à la classe supérieure quand les apprentissages importants ne sont pas

suffisamment acquis. On tient compte alors de sa motivation, de ses réactions possibles, etc.

Il y a un ensemble de précautions à prendre pour minimiser les erreurs de décision:

— travailler en comité multidisiciplinaire où plusieurs enseignants peuvent discuter de ces cas. Toutes les autres personnes qui connaissent l'élève peuvent apporter un éclairage utile: orthopédagogue, psychologue, professionnelle ou professionnel, parents. L'enseignant, l'enseignante ou le comité recommandent par la suite à la direction de l'école la décision considérée la meilleure pour l'élève. C'est alors la directrice ou le directeur de l'école qui endosse la responsabilité finale de la décision. Il va sans dire que ces décisions doivent être prises avec beaucoup de précautions et de rigueur;

— allouer suffisamment de temps pour qu'on puisse peser le pour et le contre de chacune des décisions;

— assurer une atmosphère favorable au travail et un fonctionnement qui permet à toutes les personnes de s'exprimer ouvertement — les subjectivités injustes ont alors la possibilité de résorber;

— recevoir toute demande de révision d'une décision.

Puisque l'on se situe ici à la fin de la démarche d'évaluation sommative, la qualité des décisions repose sur la qualité de la planification et de la réalisation des étapes précédentes, la mesure et le jugement.

Les responsabilités sont partagées. L'élève participe selon son niveau de développement et ses parents sont souvent impliqués. L'enseignante ou l'enseignant recommande une décision à la direction de l'école, celle-ci portant la responsabilité finale de la décision.

5.4 Comment consigner les jugements et les décisions?

Durant la démarche d'évaluation sommative, les données des diverses prises d'informations ont été consignées de façon systématique à l'aide de tableaux. Comme nous l'avons vu, certains tableaux permettent d'enregistrer les réussites aux divers objectifs ou dimensions retenus. On y inscrit aussi toute autre information pertinente qui éclairera le jugement final et la décision. Une fois les jugements portés, il s'agit de les inscrire dans le bulletin. Les décisions concernant le passage à la classe supérieure et l'octroi des crédits y sont aussi consignées.

Il existe de nombreux modèles de bulletin. Un bulletin dit traditionnel a été généralement utilisé à travers la province. Depuis quelques années toutefois, le bulletin descriptif est de plus en plus répandu*. Ceci correspond au besoin de décrire plus précisément les jugements portés sur les apprentissages et sur les comportements de l'élève. Comme le bulletin est un instrument de communication, le langage qui y est utilisé doit être clair pour que la personne qui le reçoit puisse l'interpréter correctement.

Lors de la planification de l'évaluation sommative, on a procédé au choix du bulletin ou on a pris connaissance et révisé les choix déjà faits. On a aussi prévu la façon d'utiliser le bulletin.

Dans un bulletin descriptif, on inscrit le jugement que l'on porte sur la compétence de l'élève en rapport avec les grandes parties importantes du programme.

La partie du bulletin réservée à des fins sommatives diffère de celle qui a pour fonction de communiquer des jugements qui amènent essentiellement des décisions d'aide im-

* Pour de plus amples renseignements, voir le document *Vers un bulletin scolaire renouvelé*. Gouvernement du Québec, Document d'information 16-7513.

médiate à l'élève. De plus, les renseignements qu'on donne sous forme de bilan sont moins détaillés. Les bilans communiquent le degré de réalisation des apprentissages. Ces jugements sont exprimés à l'aide d'un ensemble de symboles qui correspondent généralement à une échelle.

Quel que soit le formulaire de consignation adopté, le respect de certains principes favorisera la cohérence entre la consignation des jugements et des décisions et l'ensemble de la démarche d'évaluation sommative.

— La fonction essentielle de la consignation du jugement est la communication de la pensée de l'enseignant ou de l'enseignante sur l'état des apprentissages de l'élève en rapport avec le programme. Ainsi envisagée, la consignation des résultats d'évaluation sommative n'a pas pour fonction principale de motiver l'élève quoiqu'elle puisse effectivement avoir un impact sur sa motivation.

— Elle n'a pas pour but non plus d'indiquer le cheminement ou le progrès réalisé par l'élève à moins que le programme ne soit emplicite sous ce rapport. Elle doit renseigner à un moment donné sur le degré de maîtrise atteint à la fin d'un cheminement.

— La consignation ne traduit pas ce que pense l'élève, quelles que soient par ailleurs les capacités d'autonomie ou d'auto-évaluation développées par l'élève tout au long de son apprentissage. Elle relève de l'enseignante ou de l'enseignant, de la commission scolaire et du ministère de l'Éducation selon leurs responsabilités respectives. Ceci n'empêche évidemment pas l'enseignante ou l'enseignant de consulter, de vérifier, de se renseigner, ... bien au contraire.

— La consignation dans le bulletin doit être facile à interpréter. Une grande quantité d'informations risque de nuire plus que d'aider lorsque la cible visée est une population nombreuse et hétérogène. Par ailleurs, une trop grande concentration de l'information dans un nombre très réduit de symboles (par exemple: réussit ou échoue son secondaire ou son deuxième cycle du primaire) rend impossible leur interprétation. Même s'il s'agit de jugements globaux, ceux-ci doivent être assez nuancés pour indiquer la compétence de l'élève.

— Une interprétation utile, surtout pour les parents et pour les enseignants qui recevront l'élève dans une autre classe, est de comparer le niveau de réussite de l'élève à celui des autres. Pour chaque échelon de l'échelle adoptée, on indique en pourcentage le nombre d'élèves qui s'y retrouvent. Ainsi, à titre d'exemple, on pourrait avoir la distribution suivante: 30% des élèves maîtrisant bien l'ensemble du programme, 63% maîtrisant les exigences minimales, 5% ayant besoin de mesures d'appui et 2% en situation d'échec.

— Pour ce qui les concerne, l'enseignante et l'enseignant doivent pouvoir prouver que le jugement porté constitue un résumé exact, pertinent et significatif du niveau atteint par l'élève dans la maîtrise de l'ensemble des éléments obligatoires du programme.

— Les rapports concernant les apprentissages doivent être séparés de ceux qui se rattachent aux comportements de l'élève. L'information est alors plus claire et les règles de décision peuvent s'appliquer plus directement.

5.5 Résumé

Au moment du jugement et de la décision, on applique les règles établies lors de la planification.

Même si plusieurs façons de faire sont possibles, le jugement est toujours porté à partir d'informations pertinentes à la décision à prendre. De plus, le jugement final doit être fait le plus près possible du moment de la prise de décision. Qu'il s'agisse de porter les jugements ou de prendre les décisions, de nombreuses précautions sont à prendre pour éviter les erreurs. Dans les deux cas aussi les responsabilités sont partagées entre les divers agentes ou agents impliqués. Une fois les jugements posés ou les décisions prises, il faut être prêt à les remettre en question si les résultats ne semblent pas adéquats ou justes.

En évaluation sommative, on consigne dans le bulletin scolaire les jugements portés sur la compétence de l'élève en regard des apprentissages au programme. On y consigne également les décisions prises. Le bulletin utilisé doit favoriser, pour toutes les personnes intéressées, la compréhension des informations ainsi communiquées.

5.6 Références

GOUVERNEMENT DU QUÉBEC, Ministère de l'Éducation, Direction générale de l'évaluation, Direction du développement en évaluation. *Vers un bulletin scolaire renouvelé. Orientations du ministère de l'Éducation, consignation, transmission et conservation des résultats scolaires, Classes du primaire et du secondaire,* Québec, juillet 1985, 59 p.

NITKO, Anthony J. *Educational Tests and Measurement an Introduction,* Toronto, Harcourt Brace Jovanovich, 1983, 661 p.

WOLFS, José Luis, *Quelques propositions de notation et d'évaluation certificative dans le cadre d'une relation éducative basée sur la transparence*, Rencontre Belgo-Franco-Suisse sur l'évaluation, Genève, 3-5 septembre 1984, 15 p.

Conclusion

Dans ce fascicule, on a présenté l'évaluation sommative comme une démarche qui nécessite beaucoup de rigueur tant au moment de la planification qu'au cours de sa réalisation. On a également considéré l'évaluation sommative comme exigeant une bonne concertation entre les divers agents éducatifs impliqués. La «Politique d'évaluation pédagogique» du ministère de l'Éducation désigne la justice et l'égalité comme valeurs fondamentales lorsqu'il s'agit d'évaluation et c'est en vue d'assurer plus de justice et plus d'égalité pour chaque élève que seront consentis ces efforts de rigueur et de concertation. Même si les personnes impliquées ont des conceptions quelque peu différentes de ces valeurs, celles-ci demeurent un point de repère, un guide qui éclaire des situations autrement fort confuses et problématiques.

Dans une perspective d'amélioration des pratiques d'évaluation sommative, les défis sont grands:

— structurer l'étape jugement pour lui donner toute son importance;
— produire des instruments de mesure qui correspondent bien aux programmes d'études;
— dégager des consensus à propos de la compétence que l'on exige par rapport à chacun des programmes d'études;
— former suffisamment les divers agents de l'éducation

La recherche de la qualité en évaluation sommative est garante de la crédibilité de notre système scolaire.

Remerciements

La direction du développement de l'évaluation tient d'abord à remercier Marie-Thérèse Boisvert, rédactrice du présent fascicule et Rita-Éva Lapointe, responsable du projet.

Elle veut aussi souligner la collaboration obtenue des membres de Groupe de liaison des docimologues en milieu scolaire (G.L.D.M.S.) qui ont procédé à la conception de la série «Éléments de docimologie» et participé à sa réalisation. La Direction remercie également les organismes qui ont permis cette participation.

Pierre Bérubé
C.S.R. de Chambly

Marie-Thérèse Boisvert
C.E.C. de Montréal

Wasfi Daoud
C.S.R. Honoré-Mercier

Michel Labelle
C.S.Huntingdon

Pierre Lafleur
C.S.R. Blainville-Deux-Montagnes

Rita-Éva Lapointe
Université de Moncton

Lise Raymond-Bérubé
Consultante

François Sarrazin
C.S. Baldwin-Cartier

La Direction du développement de l'évaluation tient également à remercier toutes les personnes qui, à un moment ou à l'autre de la réalisation de ce fascicule, l'ont lu, relu, analysé et commenté, tant au plan des idées que de la forme. À cet égard, la Direction veut mentionner de façon particulière la contribution de ses propres membres. Elle souligne aussi la participation des personnes suivantes qui, répondant soit à une demande de la responsable du projet, soit à l'invitation de la personne responsable du dossier de l'évaluation dans leur Direction régionale, nous ont fait parvenir leurs commentaires sur l'avant-dernière version du présent fascicule:

Normande Beaudet
D.R. de l'Abitibi-Témiscamingue

Marcel Bédard
C.S. Le Gardeur

Marcel Bélanger
D.R. de Montréal

Louise Blais
D.R. de l'Abitibi-Témiscamingue

Gilles Boisvert
D.R. de Trois-Rivières

Pierre Boivin
C.S. Jérôme-Le Royer

Clément Brien
C.S.R. La Vérendrye

Louise Bussières-Gagné
D.R. Laval/Laurentides/Lanaudière

Andrée Chalifoux
C.S. de Brossard

Réjeanne Croteau
C.S. de Warwick

Chantal Deschamps
Université de Montréal, étudiante

Odette Filion
C.S. des Manoirs

Hélène Freidlan
Ministère de l'Éducation

Louis Gagné
Ministère de l'Éducation

Gilles Gamache
C.S. Les Écores

Diane Hinse
C.S. de Warwick

Roger Hivon
C.S.R. Meilleur

Raymond Julien
D.R. de Trois-Rivières

Gilles Lefebvre
C.S. Pierre-Neveu

Michelle LeQuellec
C.S. Laurentides

Denise Lussier
D.R. de la Montérégie

Pierre Marques
C.S. des Mille-Îles

Robert Michaud
D.R. de l'Abitibi-Témiscamingue

Claude Montpetit
C.S. Chomedey Laval

Nicole Parent-Gingras
C.S. de Saint-Jérôme

Raymond Plourde
C.S.R. de l'Yamaska

Nicole Raymond
C.S. Huntingdon

Marcel Robillard
C.S. de Soulanges

Serge Séguin
Université du Québec à Montréal

Claude Turcotte
C.S. de Saint-Jérôme

Yvon Veilleux
C.S. de Warwick

La Direction du développement de l'évaluation tient enfin à souligner l'apport de Dominique Morissette, professeur à l'université du Québec à Trois-Rivières, qui en plus d'avoir fait parvenir ses commentaires sur l'avant dernière version, a rédigé les premières versions du fascicule et participé à plusieurs rencontres dans le cadre de ce projet.

Composition typographique: GROUPE SAJY inc.

Achevé d'imprimer à Montmagny
par les travailleurs des ateliers Marquis Ltée
en mars 1988

La série *Éléments de docimologie* sera composée de cinq fascicules. Le présent ouvrage en est le quatrième. Il constitue un outil indispensable pour les personnes intéressées à l'évaluation pédagogique. Voici un aperçu du contenu de ces publications:

1 L'évaluation pédagogique: une démarche

Ce document présente une conception de l'évaluation pédagogique, distingue et illustre les trois étapes de la démarche d'évaluation (mesure, jugement, décision) et décrit deux aspects complémentaires de l'évaluation pédagogique: l'évaluation formative et l'évaluation sommative.

2 Le lexique

Non seulement existe-t-il un grand nombre de termes se référant à la mesure et à l'évaluation, encore y a-t-il différentes définitions pour chacun de ces termes. Le lexique rassemble les principaux termes utilisés dans la série «Élément de docimologie» et les définit selon le cadre théorique présenté dans le fascicule portant sur l'évaluation pédagogique.

3 L'évaluation formative

Ce fascicule clarifie les concepts et présente de façon détaillée les étapes à réaliser dans une démarche d'évaluation formative. On y trouve de nombreux exemples concrétisant cette démarche qui vise à favoriser la progression constante des apprentissages.

4 L'évaluation sommative

Ce fascicule clarifie les concepts et présente de façon détaillée les étapes à suivre pour compléter une démarche d'évaluation sommative. On y trouve quelques exemples illustrant cette démarche qui vise à faire le bilan des apprentissages réalisés par l'élève.

5 Les instruments de mesure (à paraître)

Il existe plusieurs types d'instruments de mesure et d'appréciation. Ce fascicule les présente tour à tour, fournissant les directives et les conseils les plus récents pour les construire, les utiliser et pour en interpréter les résultats.

Gouvernement du Québec
Ministère de l'Éducation

9 782551 082070

Imprimé au Québec, Canada